THREE ESSAYS IN MICROECONOMICS

BY

KANG RONG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Steven Williams, Chair
Professor Dan Bernhardt
Professor Bart Taub, Durham University
Professor Mattias Polborn

Abstract

This dissertation consists of three essays in microeconomics. The first essay studies a finite-horizon alternating-offer model that integrates the common practice of having an arbitrator determine the outcomes if both players' offers are rejected. We find that if the arbitrator uses final-offer arbitration (as in professional baseball), and the arbitrator does not excessively favor one player, then the unique subgame-perfect equilibrium always coincides with the subgame-perfect equilibrium outcome in Rubinstein's infinite-horizon alternating-offer game. However, if the arbitrator sufficiently favors the player making the initial offer, then delay occurs in equilibrium. If, instead, the arbitrator uses the split-the-difference arbitration rule, then the unique subgame-perfect equilibrium can feature immediate agreement, delayed agreement, or no agreement, depending on the discount factor.

The second essay studies the arbitration problem using the axiomatic approach. In particular, we define an arbitration problem as the triplet of a bargaining set and the offers submitted by two players. We characterize the solution to a class of arbitration problems using the axiomatic approach. The axioms we impose on the arbitration solution are "Symmetry in Offers," "Invariance" and "Pareto Optimality." The key axiom, "Symmetry in Offers," requires that whenever players' offers are symmetric, the arbitrated outcome should also be symmetric. We find that there exists a unique arbitration solution, called the *symmetric arbitration solution*, that satisfies all three axioms. We then analyze a simultaneous-offer game and an alternating-offer game. In both games, the symmetric arbitration solution is used to decide the outcome whenever players cannot reach agreement by themselves.

We find that in both games, if the discount factor of players is close to 1, then the unique subgame perfect equilibrium outcome coincides with the Kalai-Smorodinsky solution outcome.

The third essay studies the public good provision problem in which a public good can be provided and payments can be collected from agents only if the proportion of agents who obtain nonnegative expected utilities from the public good provision mechanism weakly exceeds a prespecified ratio α. We call this requirement "α proportional individual rationality". We identify a key threshold such that if α is less than this threshold, then efficiency obtains asymptotically. If α is greater than the threshold, then inefficiency obtains asymptotically. In addition, we obtain the convergence rate of the probability of provision to its efficient/inefficient level.

To My Wife, My Parents, and My Sister

Acknowledgements

I thank Steven Williams, my advisor, for his patient guidance and helpful advice. I would also like to thank my committee members, Dan Bernhardt, Bart Taub and Mattias Polborn, for their valuable comments and kind support.

Table of Contents

Chapter 1

Finite-Horizon Alternating-Offer Game with Arbitration

1.1 Introduction

The industrial relations literature features two types of well-known arbitration procedures: conventional arbitration, and final-offer arbitration. Conventional arbitration is an arbitration process in which the arbitrator is free to impose any settlement as the arbitration outcome. A simple, widely-used conventional arbitration procedure is (equally) splitting the difference between players' final offers (e.g., Anbarci and Boyd 2011; Compte and Jehiel 1995). In contrast, in final-offer arbitration, the arbitrator must choose one player's final offer as the arbitration outcome. Final-offer arbitration was first proposed by Stevens (1966) and has been widely used in the public-sector to settle labor disputes and in major league baseball to resolve salary disputes (Chelius and Dworkin 1980; Wilson 1994).

Although arbitration is a common dispute resolution mechanism, it has received little attention in the bargaining literature. The purpose of my paper is to explore how introducing arbitration affects players' equilibrium strategies and bargaining outcomes. My paper will address the following questions: when does the introduction of arbitration have an impact on the equilibrium of the bargaining game? and if arbitration has an impact, what kind of impact will it have and which player benefits? I show that if final-offer arbitration is used, then as long as the arbitrator does not excessively favor one player, the equilibrium of the game is *unaffected* by the specific details of the arbitrator's preference, and both players obtain Rubinstein equilibrium payoffs. In all other cases where final-offer arbitration is used, the equilibrium of the game depends on the arbitrator's preference and the player favored by the arbitrator obtains a payoff higher than his Rubinstein equilibrium payoff. In addition, I show that delay in reaching agreement emerges when the arbitrator is excessively biased toward the player who makes the first offer: the bias encourages the player making the first offer to make an unattractive demanding offer in order to get "closer" to the threat of having allocations determined by the biased arbitrator.

If, instead, the split-the-difference arbitration rule is used, then the equilibrium of the game depends on the discount factor. In particular, when the discount factor is sufficiently small, the equilibrium features immediate agreement. When the discount factor is sufficiently large, the equilibrium is an equilibrium with no agreement, where both players make extreme demands and both offers are rejected (so the arbitration outcome is the final outcome). Delay emerges when players are sufficiently patient: players make extreme demands in equilibrium because the time cost of going to arbitration is small and when arbitration is reached, split-the-difference punishes compromise offers.

My basic framework builds on Yildiz (2011). Two players sequentially make offers. If a player's offer is accepted by his opponent, then that offer is the bargaining outcome and the game ends. If, instead, both players' offers are rejected by their opponents, an arbitration stage is reached. In contrast to Yildiz (2011), who assumes that the arbitrator chooses the offer that yields the higher Nash product as the arbitration outcome, I consider two very general classes of arbitration rules. One class is the family of final-offer arbitration rules, where the arbitrator's *ideal settlement* can be any point on the Pareto frontier of the bargaining set and the arbitrator chooses his preferred offer, i.e., the offer closest to his ideal settlement as the arbitration outcome. The other class is the split-the-difference arbitration rule.

Player 1 makes the first offer. When final-offer arbitration is used, I find that, (i) if the arbitrator is "balanced" (i.e., the arbitrator does not excessively favor one player),[1] then the unique subgame perfect equilibrium (henceforth SPE) outcome of the game coincides with the equilibrium outcome of Rubinstein's infinite-horizon alternating-offer game (Rubinstein 1982); (ii) if the arbitrator sufficiently favors Player 1, then the unique SPE of the game is such that Player 1 makes an offer that Player 2 *rejects*, and Player 2 makes a counteroffer that Player 1 accepts; the equilibrium payoff received by Player 1 exceeds what he would obtain from the Rubinstein equilibrium; and (iii) if the arbitrator sufficiently favors Player 2, then the unique SPE of the game is such that Player 1 makes a more generous offer than the Rubinstein equilibrium offer that Player 2 accepts immediately.

In the game that Yildiz (2011) considers, the unique SPE outcome coincides with the Rubinstein equilibrium outcome. This result might lead one to believe that the arbitrator's preference used by Yildiz (2011) is special to the extent that it may

[1] An arbitrator is "balanced" if the arbitrator's ideal settlement is close to the Rubinstein equilibrium outcome. The measure of closeness depends on the discount factor.

be the only arbitrator's preference for which the SPE outcome coincides with the Rubinstein equilibrium outcome. However, my analysis shows that there exists a broad class (depending on the discount factor) of arbitrator's preferences for which the unique SPE outcome coincides with the Rubinstein equilibrium outcome. What is special about the arbitrator's preference that Yildiz (2011) considers is that it belongs to that class for *all* discount factors.

One implication of my analysis is the following *irrelevance* result: as long as the arbitrator is not too biased toward a player, then the unique equilibrium of the arbitration game is unaffected by the arbitrator's preference. In reality, when the arbitrator uses final-offer arbitration, people may have concerns about the fairness of the arbitrator. However, my irrelevance result shows that outcomes are unaffected if the arbitrator has some bias, as long as this bias is not too large. In other words, there is a wide range of arbitrator's preferences under which the equilibrium of the arbitration game is independent of the arbitrator's preference. Within this range, the precise choice of the arbitrator becomes unimportant.

Another implication of my analysis is that even when players have complete information, delay might arise due to the introduction of arbitration.[2] Delay in equilibrium occurs when the arbitrator sufficiently favors Player 1. The intuition is as follows. If Player 1 demands more than the Rubinstein equilibrium outcome, then Player 2 will reject Player 1's offer to exploit "time delay". That is, Player 2 will make a counteroffer, which Player 1 accepts in order to avoid the time cost of going

[2]Delay in equilibrium within the framework of complete information also occurs in Manzini and Mariotti (2001), Manzini and Mariotti (2004), Ponsatí and Sákovics (1998) and Rubinstein (1982). However, the mechanism for delay is different. It arises in those models due to the existence of multiple equilibria.

to arbitration, and Player 2 is better off by making such a counteroffer, instead of accepting Player 1's initial offer.[3] If, instead, Player 1 demands less than the Rubinstein equilibrium outcome, then the offer is accepted by Player 2 immediately. If the arbitrator *sufficiently* favors Player 1 (for a given discount factor), then Player 1 prefers to demand more than the Rubinstein equilibrium outcome. This is because Player 1 can demand a payoff that is close to the arbitrator's ideal settlement, which is *sufficiently* higher than Player 1's Rubinstein equilibrium payoff. Such a demand is supported by the threat of the biased arbitrator, which implies that Player 2's counteroffer cannot be far away from Player 1's offer. As a result, the benefit that Player 1 can exploit from the biased arbitrator exceeds the cost incurred from the delayed agreement. Therefore, delay in equilibrium occurs.

Manzini and Mariotti (2001) considered an infinite-horizon alternating-offer model that also involves arbitration. They assume that an arbitrator can be called in whenever a player has just rejected an offer and both players agree to move to arbitration. My result contrasts with theirs in the sense that they show that the Rubinstein equilibrium can arise only if the arbitration outcome sufficiently *favors* one of the players, while my result shows that the Rubinstein equilibrium arises only if the arbitration rule does *not excessively favor* a player. The difference between the results is due to the following: (i) In Manzini and Mariotti (2001), the alternating-offer game has an *infinite horizon* and the arbitration outcome is *exogenously* given

[3] In particular, assume that Player 1 makes the offer $(x_1, f(x_1))$, where x_1 is Player 1's own demand and $f(\cdot)$ is the Pareto frontier function. Then, rather than accept Player 1's offer, Player 2 can always obtain a higher payoff by rejecting Player 1's offer and making the counteroffer $(\delta x_1, f(\delta x_1))$ (depending on the arbitrator's preference, Player 2 may make a more ungenerous counteroffer in equilibrium), which Player 1 will accept. Notice that $\delta f(\delta x_1) > f(x_1)$, as long as x_1 is greater than Player 1's Rubinstein equilibrium payoff.

when the arbitrator is called in. As a result, if the arbitrator is too biased toward one player, the arbitration (as a background threat) becomes a non-binding threat because both players need to agree in the event that an arbitrator is called in. Therefore, the alternating-offer game in Manzini and Mariotti (2001) yields the same equilibrium outcome as Rubinstein's alternating-offer game. (ii) In my model, the alternating-offer game has a *finite horizon* and the arbitration outcome is *endogenous* in the sense that it depends on both players' offers. As a result, when the final-offer arbitration rule is sufficiently "balanced" in the sense that the arbitrator's ideal settlement is sufficiently close to the Rubinstein equilibrium outcome, it is not profitable for Player 1 to demand more than the Rubinstein equilibrium outcome because of the time cost of delay,[4] and the equilibrium outcome of the alternating-offer game coincides with the Rubinstein equilibrium outcome.

Finally, I find that if the split-the-difference arbitration rule is used, then the unique SPE depends on the common discount factor of players. In particular, (i) if the discount factor is sufficiently small, then the unique SPE is such that Player 1 makes an offer that Player 2 accepts immediately. (ii) If the discount factor is sufficiently large, then the unique SPE is such that both players make extreme offers and both offers are rejected by the opponents (so the game moves on to the arbitration stage). This result is due to the chilling effect of conventional arbitration. That is, as players become more patient, they are more likely to take extreme positions before

[4] More particularly, for Player 1, making an excessively high demand is not supported by the arbitrator and is thus not profitable. Making a demand that is only slightly higher than the Rubinstein equilibrium offer is also not profitable because of the time cost of delay (notice that even if Player 1 makes a demand that is slightly higher than the Rubinstein equilibrium offer, it will be rejected by Player 2; see also footnote 3).

arbitration.

This paper is organized as follows. Section 1.2 introduces some notation that helps define the "alternating-offer arbitration game". Section 1.3 studies the arbitration game that uses the final-offer arbitration rule. Section 1.4 studies the arbitration game that uses the split-the-difference arbitration rule. Concluding remarks are offered in Section 1.5.

1.2 The Model

There are two players, Players 1 and 2, who are expected utility maximizers. Let $S \subset R^2$ denote the bargaining set, which includes all possible bargaining outcomes, measured in terms of expected utility level. I use $(x_1, y_1) \in S$ to denote Player 1's offer and use $(x_2, y_2) \in S$ to denote Player 2's offer, where x represents Player 1's utility payoff and y represents Player 2's utility payoff. I normalize the disagreement point of S to $(0, 0)$. I assume that $(x, y) \geq (0, 0)$ for any $(x, y) \in S$, and that there is at least one point $(x, y) \in S$ such that $(x, y) >> (0, 0)$. The bargaining set S is assumed to be convex, compact and strictly comprehensive. The bargaining set S is *comprehensive* if $(x', y') \in S$ whenever $(0, 0) \leq (x', y') \leq (x, y)$ and $(x, y) \in S$. The bargaining set S is *strictly comprehensive* if S is comprehensive and there exists a $(x'', y'') \in S$ such that $(x'', y'') >> (x, y)$ whenever $(x, y) \in S$ and $(x', y') \in S$ with $(x', y') \geq (x, y)$ and $(x', y') \neq (x, y)$.

The (weak) Pareto frontier of the bargaining set S is defined as $PF = \{p \in S : q >> p \Rightarrow q \notin S\}$.[5] Define $b_i = max\{U_i : (U_1, U_2) \in S\}$ to be Player i's maximal

[5] PF depends on S. However, I fix S in this paper, and I omit the dependency on S in notation whenever there is no confusion.

possible utility payoff from the bargaining set. Define $f : x \to max\{y|(x,y) \in S\}$ for $x \in [0, b_1]$. The function f is well-defined because S is compact. In addition, the assumption that S is convex and strictly comprehensive implies that f is concave, continuous, and strictly decreasing on $[0, b_1]$ with $f(0) = b_2$ and $f(b_1) = 0$ (see Figure 1). Note that $(x, y) \in PF$ if and only if $y = f(x)$.

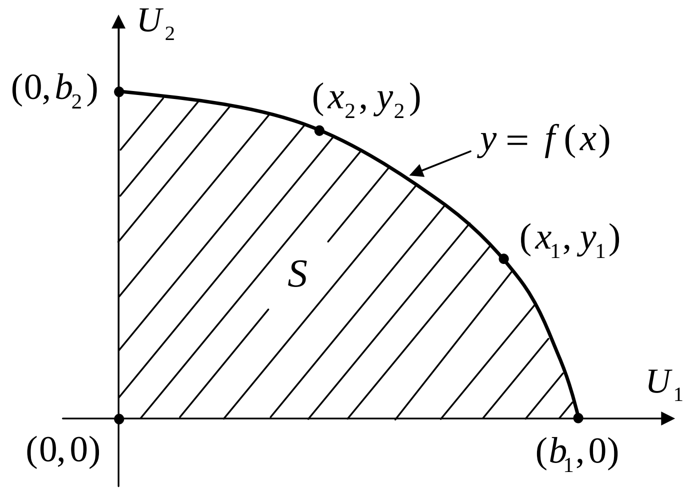

Figure 1: The bargaining set.

I assume that there is an arbitrator who is informed about players' utilities.[6] Define $\mathcal{B} = \{((x_1, y_1), (x_2, y_2)) \,|(x_1, y_1) \in S\,,\ (x_2, y_2) \in S\}$. An *arbitration rule* is a function $h : \mathcal{B} \to S$. I write $h((x_1, y_1), (x_2, y_2)) = (h_1((x_1, y_1), (x_2, y_2)), h_2((x_1, y_1), (x_2, y_2)))$, where $h_i((x_1, y_1), (x_2, y_2))$ is the arbitration outcome of Player i. A *final-offer arbitration rule* is any arbitration rule where $h((x_1, y_1), (x_2, y_2)) \in \{(x_1, y_1), (x_2, y_2)\}$ for any $((x_1, y_1), (x_2, y_2)) \in \mathcal{B}$. The *split-the-difference* arbitration rule is the rule $h((x_1, y_1), (x_2, y_2)) = (\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2})$ for any

[6]This is a standard assumption in the law and economics literature (Deck and Farmer 2007).

$((x_1, y_1), (x_2, y_2)) \in \mathcal{B}$.[7] I assume that h is known to both players.

The only arbitration cost takes the form of *time costs*, which are measured by the common discount factor δ of players. I assume that $\delta \in (0, 1)$. The unique intersection point of the curve $y = \delta f(x)$ and the curve $y = f(\frac{1}{\delta}x)$ is denoted by $(\delta x^R(\delta), f(x^R(\delta)))$. That is, $f(x^R(\delta)) = \delta f(\delta x^R(\delta))$ (see Figure 2). The point $(x^R(\delta), f(x^R(\delta)))$ is the outcome of the unique SPE of Rubinstein's infinite-horizon alternating-offer game (Rubinstein 1982).[8] Since δ is fixed in most parts of the paper, I write $(x^R(\delta), f(x^R(\delta)))$ as $(x^R, f(x^R))$ whenever it does not create confusion.

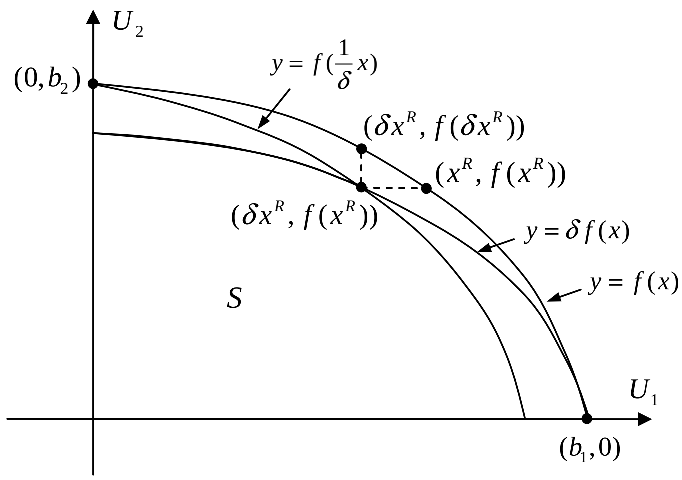

Figure 2: Definition of $(x^R, f(x^R))$.

I define the *alternating-offer arbitration game* (or simply, the *arbitration game*), which generalizes the game considered in Yildiz (2011), as the following three-stage

---

[7]Given that S is a convex set, $(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}) \in S$ for any $(x_1, y_1) \in S$ and $(x_2, y_2) \in S$.

[8]In Rubinstein's model, if Player 1 makes the Rubinstein equilibrium offer $(x^R(\delta), f(x^R(\delta)))$, then Player 2 is indifferent between accepting the offer and rejecting the offer. This is because if Player 2 accepts the offer, then his payoff is $f(x^R(\delta))$; if Player 2 rejects the offer, then at the next stage Player 2 will make the offer $(\delta x^R(\delta), f(\delta x^R(\delta)))$ in equilibrium, which Player 1 will accept, giving Player 1 a discounted payoff of $\delta f(\delta x^R(\delta)) = f(x^R(\delta))$.

procedure:

Stage 1: Player 1 makes an offer $(x_1, y_1) \in S$. Player 2 decides whether to accept the offer, ending the game with (x_1, y_1), or to reject the offer, moving the game to the next stage;

Stage 2: Player 2 makes an offer $(x_2, y_2) \in S$. Player 1 decides whether to accept the offer, ending the game with (x_2, y_2), or to reject the offer, moving the game to the arbitration stage;

Stage 3: An arbitrator decides the final outcome using arbitration rule h, i.e., $h((x_1, y_1), (x_2, y_2))$ is the arbitrated outcome.

Section 1.3 analyzes the final-offer arbitration game and Section 1.4 analyzes the split-the-difference arbitration game.

1.3 Final-Offer Arbitration Game

1.3.1 Notation and Assumptions

In this subsection, I introduce notation and assumptions regarding the final-offer arbitration rule h.

For any $(x_1, y_1) \in S$, define the set $V(x_1, y_1) = \{(x_2, y_2) \in S | h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)\}$ and $V_x(x_1, y_1) = \{x_2 \in [0, b_1] | (x_2, y_2) \in V(x_1, y_1)\}$. Thus, $V(x_1, y_1)$ is the collection of (x_2, y_2) that is chosen as the arbitration outcome when the final offers of the two players are (x_1, y_1) and (x_2, y_2), and $V_x(x_1, y_1)$ is the set of Player 1's payoffs in the set $V(x_1, y_1)$. The set $V_x(x_1, y_1)$ is nonempty because $(x_1, y_1) \in V(x_1, y_1)$ and $x_1 \in V_x(x_1, y_1)$. The set $V_x(x_1, y_1)$ is bounded because S is bounded. For any $(x_1, f(x_1))$, define $g(x_1) = \min\{x_2 | x_2 \in V_x(x_1, f(x_1))\}$. Note that $g(x_1)$ is well defined if $V_x(x_1, f(x_1))$ is closed (Condition 1 below).

Figure 15 illustrates the sets $V(x_1, f(x_1))$ and $V_x(x_1, f(x_1))$ and the function $g(x_1)$ for the case of *Nash final-offer arbitration rule.* The Nash final-offer arbitration rule is the final-offer arbitration rule such that, among the final offers submitted by the two players, the offer that yields the higher Nash product is chosen as the arbitration outcome. This is the rule that Yildiz (2011) considers.

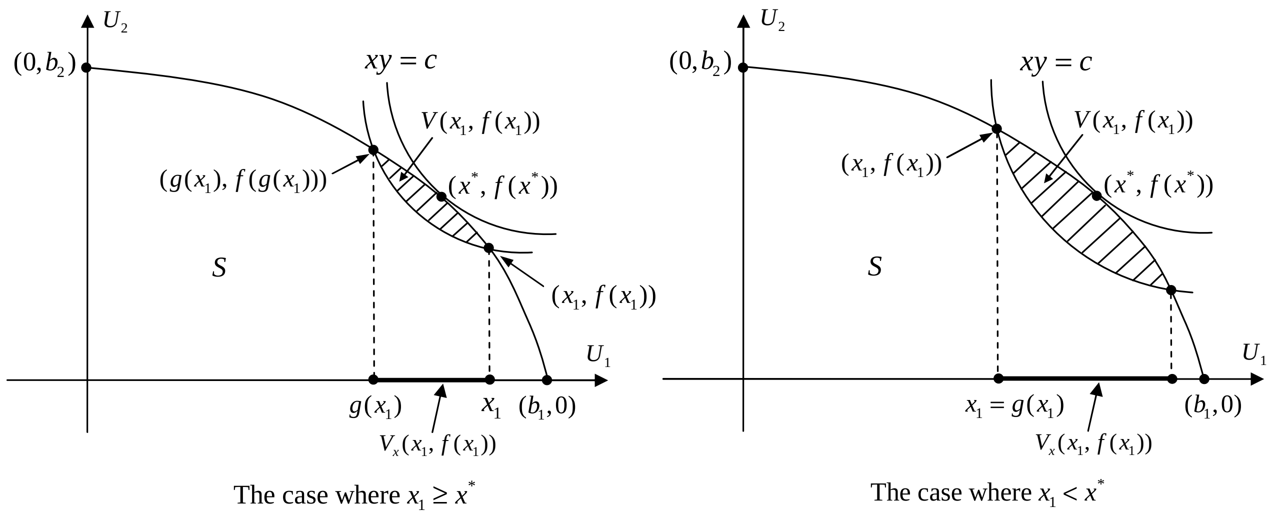

Figure 3: $V(x_1, f(x_1))$, $V_x(x_1, f(x_1))$ and $g(x_1)$ under the Nash final-offer arbitration rule.

A final-offer arbitration rule h is *regular* if there exists a continuous, strongly monotone, and quasiconcave function $u : S \to R$ such that $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ if and only if $u(x_2, y_2) \geq u(x_1, y_1)$.[9] The function u can be regarded as the arbitrator's utility function. It can be easily verified that if the final-offer arbitration

[9]I am indebted to an anonymous referee for suggesting this definition. Strong monotonicity of u requires that $u(x', y') > u(x, y)$ whenever $(x', y') \geq (x, y)$ and $(x', y') \neq (x, y)$. Quasiconcavity of u requires that the set $\{(x, y) \in S : u(x, y) \geq \overline{u}\}$ be convex for any $\overline{u} \in R$.

rule h is regular, then the set $V_x(x_1, f(x_1))$ and the function $g(x_1)$ satisfy the following three properties:

Condition 1 (Closedness). *The set $V_x(x_1, f(x_1))$ is a closed interval for any $x_1 \in [0, b_1]$.*

Condition 2 (Continuity). *The function $g(x_1)$ is continuous in x_1 for $x_1 \in [0, b_1]$.*

Condition 3 (Relative Fairness). *There exists an $x^* \in [0, b_1]$ such that, (i) for $x_1 \in [0, x^*]$, we have $g(x_1) = x_1$; (ii) the function $g(x_1)$ is decreasing on $(x^*, b_1]$.*

Condition 1 ensures that the function g is well-defined. Condition 2 ensures that the arbitration curve (defined below) is continuous.

The point $(x^*, f(x^*))$ in Condition 3 can be regarded as the arbitrator's *ideal settlement*. Roughly speaking, Condition 3 implies that (i) if Player 1's offer is *generous* to Player 2 (i.e., $x_1 \leq x^*$), then Player 2's offer will not be chosen as the arbitration outcome if Player 2 demands more than the amount that Player 1 offers him; (ii) if Player 1's offer is *ungenerous* to Player 2 (i.e., $x_1 > x^*$), then the arbitrator will allow Player 2 to make an offer where Player 2's own demand exceeds the ideal settlement, and Player 2's offer is still chosen as the arbitration outcome. Moreover, this "tolerance" for Player 2's high demand increases as Player 1's offer becomes more ungenerous.

One can show that the Nash final-offer arbitration rule is regular. Under the Nash final-offer arbitration rule, the utility function that the arbitrator maximizes is $u(x, y) = xy$, and the arbitrator's ideal settlement is the Nash bargaining solution outcome.

Let Σ denote the set of all *regular* final-offer arbitration rules. The remainder of this section focuses on regular final-offer arbitration rules.

I now define two curves. The *arbitration curve* is defined as the curve $x = g(f^{-1}(y))$ where $y \in [0, f(x^*)]$ (see Figure 4). According to Condition 3 (ii) and using the fact that f is strictly decreasing, it follows that the arbitration curve is increasing on $[0, f(x^*)]$.

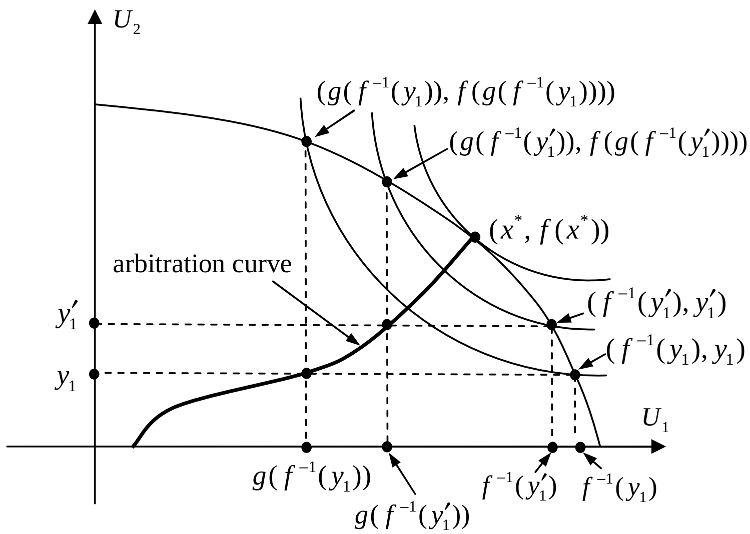

Figure 4: The arbitration curve.

The other curve, the *discounted Pareto frontier*, is defined as follows (see Figure 5):

$$y = \begin{cases} \delta f(x) & if \quad 0 \leq x \leq \delta x^R \\ f(\frac{1}{\delta}x) & if \quad \delta x^R < x \leq \delta b_1. \end{cases}$$

If the arbitration curve and the discounted Pareto frontier intersect, then there must be a unique intersection point. Denote this point by $(\hat{x}(\delta), \hat{y}(\delta))$ (see Figure 5).

In the remainder of the paper, I write $(\hat{x}(\delta), \hat{y}(\delta))$ as $(\hat{x}, \hat{y})$ whenever δ is fixed. If there is no intersection point between those two curves, then the arbitration curve must intersect the X-axis at a point that is to the right of the point $(\delta b_1, 0)$ (using the fact that the arbitration curve is continuous). In this case, define $(\hat{x}, \hat{y}) = (\delta b_1, 0)$.

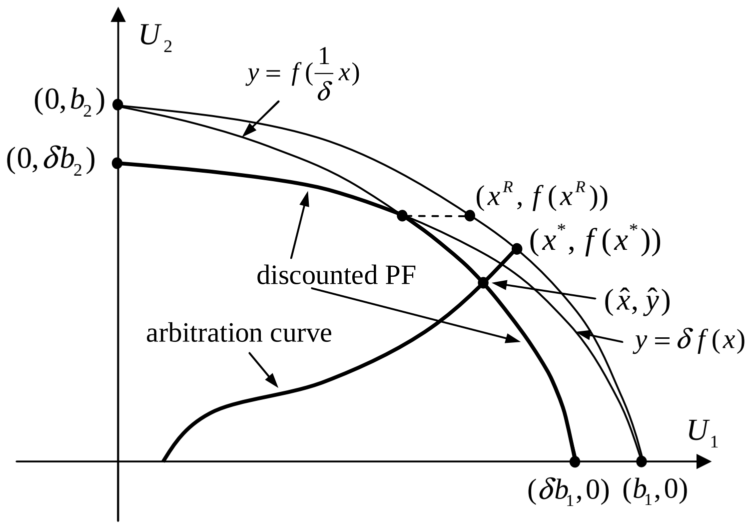

Figure 5: Definition of $(\hat{x}, \hat{y})$.

I show in Theorem 1 that the relationship between $\hat{x}$ and x^R is the key to identifying the SPE of any arbitration game that uses a regular final-offer arbitration rule.

1.3.2 Characterization of the Equilibrium

I make the following two tie-breaking rules to simplify the analysis.

Tie-breaking rule 1: If a player is indifferent between acceptance and rejection, he accepts.

Tie-breaking rule 2: If a player is indifferent between two options that he can offer his opponent, he chooses the option that yields his opponent a higher payoff.

The following result characterizes the SPE of the arbitration game that uses arbitration rule $h \in \Sigma$.

Theorem 1. *In the arbitration game with arbitration rule $h \in \Sigma$, we have:*

(i) (Rubinstein equilibrium.) If $\delta x^R \leq \hat{x} \leq \frac{1}{\delta} x^R$, then the outcome of the unique SPE is that Player 1 makes the offer $(x^R, f(x^R))$ and Player 2 accepts it;

(ii) (type-II arbitration-driven equilibrium.) If $\frac{1}{\delta} x^R < \hat{x} \leq \delta b_1$, then the outcome of the unique SPE is that at Stage 1, Player 1 makes the offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$, which Player 2 rejects, and at Stage 2, Player 2 makes the offer $(\hat{x}, f(\hat{x}))$, which Player 1 accepts;

(iii) (type-I arbitration-driven equilibrium.) If $0 \leq \hat{x} < \delta x^R$, then the outcome of the unique SPE is that Player 1 makes the offer $(f^{-1}(\delta f(\hat{x})), \delta f(\hat{x}))$ and Player 2 accepts it.

Proof: See Appendix 1. □

A final-offer arbitration rule h is *balanced* (or, an arbitrator is *balanced*) if $\hat{x} \in [\delta x^R, \frac{1}{\delta} x^R]$.[10] Roughly speaking, the balancedness of an arbitration rule requires that the arbitrator's ideal settlement be close enough to the Rubinstein equilibrium outcome.[11] According to Theorem 1 (i), if a regular final-offer arbitration rule is

[10]The definition of "balanced" is distinct from the definition of "balanced" in Manzini and Mariotti (2001). In both papers, an arbitration rule is balanced if and only if the arbitration rule does not excessively favor a player. However, the measure of "favorness" is different.

[11]This is because the arbitrator's ideal settlement is close to $(\hat{x}, \hat{y})$, and $(\hat{x}, \hat{y})$ is close to the Rubinstein equilibrium outcome by balancedness. The measure of "closeness" depends on the discount factor.

balanced, then the equilibrium outcome is such that Player 1 offers $(x^R, f(x^R))$, which Player 2 accepts. This result follows from the following two facts. First, one can show that Player 1's offer is accepted by Player 2 if and only if Player 1's demand is less than the Rubinstein equilibrium payoff (Lemma 11 (i) in Appendix 1), so the best offer that Player 1 can make and Player 2 accepts is $(x^R, f(x^R))$. Second, if Player 1 makes a demand that is higher than the Rubinstein equilibrium offer, then Player 1's offer is rejected by Player 2, and Player 2 makes a counteroffer that Player 1 accepts (Lemma 11 (i) and Lemma 10 in Appendix 1). When the final-offer arbitration rule is balanced in the sense that the arbitrator's ideal settlement is close enough to the Rubinstein equilibrium outcome, Player 2's counteroffer is at most marginally more generous than the Rubinstein equilibrium outcome. For Player 1, the extra benefit of making an offer that will be rejected by Player 2 is thus less than the time cost incurred by reaching a delayed agreement. As a result, when the arbitrator is balanced, Player 1 makes the offer $(x^R, f(x^R))$ that Player 2 accepts immediately.

For the class of balanced final-offer arbitration rules, the details of the final-offer arbitration rule are irrelevant to the equilibrium outcome. That is, letting $\Sigma^* = \{h | h \in \Sigma$ with the corresponding $\hat{x} \in [\delta x^R, \frac{1}{\delta} x^R]\}$, we have:

Corollary 2. *(Irrelevance Result) For any $h \in \Sigma^*$ and $h' \in \Sigma^*$, the arbitration game that uses rule h yields the same equilibrium outcome as the arbitration game that uses rule h'.*

To understand the significance of the above irrelevance result, we can imagine that there is a stage before the arbitration game. At this pre-arbitration stage, an arbitrator is chosen from a pool, in which different arbitrators may have different preferences over the two players' offers. My irrelevance result shows that there is a pool of arbitrators, in which the choice of the arbitrator does not matter for the equilibrium of the arbitration game. Moreover, this pool of arbitrators is reasonable and sufficiently wide in the sense that it includes all arbitrators who are *not too biased* toward a player. However, as the discount factor increases, this pool shrinks. When the discount factor approaches 1, the only arbitrator that belongs to this pool *must* be the arbitrator whose ideal settlement is the Nash bargaining solution outcome (see Appendix 4.1 for the robustness of the Nash final-offer arbitration rule).

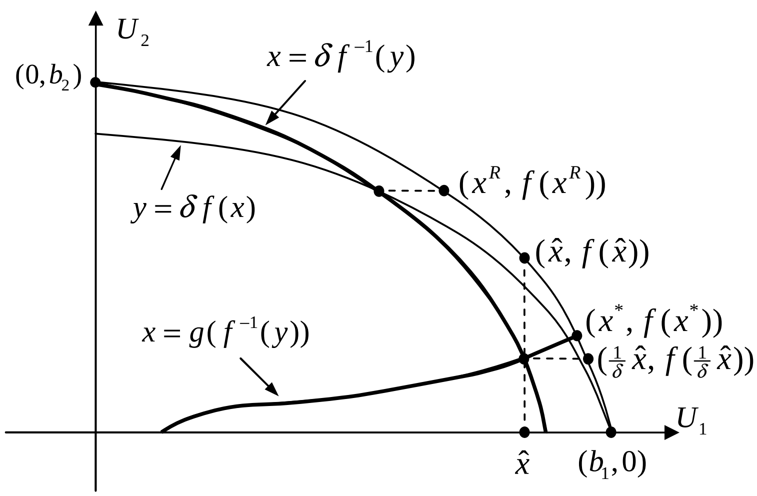

Figure 6: The equilibrium with delay.

If $\frac{1}{\delta}x^R < \hat{x} \leq \delta b_1$, then the unique SPE is a type-II arbitration-driven equilibrium, which is an equilibrium with delayed agreement. One may wonder why it is optimal for Player 1 to make an offer that will be rejected by Player 2, as opposed to making an offer that will be accepted by Player 2. The reason is as follows. On one hand, if Player 1 makes an offer that demands more than the Rubinstein equilibrium payoff, then the offer will be rejected by Player 2, and Player 2 will make a counteroffer that Player 1 accepts (Lemma 11 (i) and Lemma 9 in Appendix 1). However, if the arbitrator sufficiently favors Player 1 ($\frac{1}{\delta}x^R < \hat{x} \leq \delta b_1$), then Player 2's counteroffer will be close enough to Player 1's initial offer, as long as Player 1's demand is not excessively higher than the arbitrator's ideal settlement. One can show that if Player 1 makes the offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$, then Player 2's counteroffer $(\hat{x}, f(\hat{x}))$ is the most favorable counteroffer that Player 1 could possibly obtain (see Figure 6).[12] On the other hand, if Player 1 makes an offer that demands less than the Rubinstein equilibrium payoff, then the offer will be accepted by Player 2 (Lemma 11 (i) in Appendix 1). Thus, the maximum payoff that Player 1 can obtain by making an offer that will be accepted by Player 2 is the Rubinstein equilibrium payoff. Since $\delta\hat{x} > x^R$, it is optimal for Player 1 to make an offer that will be rejected by Player 2. Therefore, delay in equilibrium occurs.

Notice that even if Player 1 makes the offer $(x^*, f(x^*))$, which is the arbitrator's ideal settlement, Player 2's equilibrium action is to reject Player 1's offer and make the counteroffer $(\delta x^*, f(\delta x^*))$. Moreover, Player 1 will accept the counteroffer in order to avoid the time cost of going to arbitration. However, in equilibrium, Player 1 will not make the offer $(x^*, f(x^*))$, because he can obtain a more favorable counteroffer by making the offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$ since $\hat{x} > \delta x^*$.

Finally, if the final-offer arbitration rule is sufficiently biased in favor of Player 2,

[12] The reason is that if Player 2 rejects Player 1's offer (x_1, y_1) at Stage 1, then at the next stage, Player 2 makes the counteroffer $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$, which Player 1 will accept (Lemma 9 in Appendix 1). Thus, Player 1 can "control" Player 2's optimal counteroffer by varying x_1. The most favorable counteroffer to Player 1 occurs when x_1 is such that $\delta x_1 = g(x_1)$, i.e., $x_1 = \frac{1}{\delta}\hat{x}$ (see Figure 6).

then the equilibrium is an equilibrium with immediate agreement. In addition, Player 1's equilibrium offer is more generous than the Rubinstein equilibrium offer. This is because if Player 1 makes an offer that is *not* more generous than the Rubinstein equilibrium offer, then Player 2 will reject the offer and make a counteroffer, in which Player 2's demand is sufficiently higher than the Rubinstein equilibrium offer. Such a counteroffer is supported by the biased arbitrator, and Player 1 has to accept it. Player 1 is thus better off to make a more generous offer than the Rubinstein equilibrium offer, which Player 2 accepts immediately.

1.3.3 Kalai-Smorodinsky Final-Offer Arbitration

This section studies the final-offer arbitration rule in which the arbitrator's ideal settlement is the Kalai-Smorodinsky solution outcome.

Definition 3. *(Kalai and Smorodinsly 1975) The Kalai-Smorodinsky (KS) solution outcome* $(x^{KS}, f(x^{KS}))$ *is the intersection point of the Pareto frontier with the line connecting* $(0,0)$ *and* (b_1, b_2)*, i.e.,* $\dfrac{f(x^{KS})}{x^{KS}} = \dfrac{b_2}{b_1}$.

Suppose the arbitrator's utility function is u^{KS}, where $u^{KS} : S \to R$ is a continuous, strongly monotone, and quasiconcave function with

$$u^{KS}(x,y) = \begin{cases} (\dfrac{y/x}{b_2/b_1})^{-1} & if \quad \dfrac{y/x}{b_2/b_1} \geq 1; \\ \dfrac{y/x}{b_2/b_1} & if \quad \dfrac{y/x}{b_2/b_1} < 1. \end{cases}$$

for $(x, y) \in PF$.[13] Notice that u^{KS} is maximized at the Kalai-Smorodinsky solution outcome. We have:

[13] I do not restrict the function u^{KS} for (x, y) inside the bargaining set, as long as u^{KS} is con-

Theorem 4. *Suppose the final-offer arbitration rule is such that* $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ *if and only if* $u^{KS}(x_2, y_2) \geq u^{KS}(x_1, y_1)$*. Then, as long as* $\frac{f(\frac{1}{\delta}x^R)}{\frac{1}{\delta}x^R} \leq \frac{b_2}{b_1} \leq \frac{f(x^R)}{\delta x^R}$*, the unique SPE outcome of the arbitration game is* $(x^R, f(x^R))$*.*

Proof: See Appendix 4.2. □

Theorem 4 implies that as long as the line that connects the origin to (b_1, b_2) is above the line that connects the origin to $(\frac{1}{\delta}x^R, f(\frac{1}{\delta}x^R))$ and is below the line that connects the origin to $(\delta x^R, f(x^R))$, then the unique SPE outcome of the alternating-offer arbitration game is $(x^R, f(x^R))$ (see Figure 4).

The condition in Theorem 4 is a sufficient condition for the Rubinstein equilibrium outcome in the KS final-offer arbitration game. More generally, Appendix 4.2 provides the sufficient and necessary condition for the Rubinstein equilibrium outcome and the conditions for other types of equilibrium (type-I arbitration-driven equilibrium and type-II arbitration-driven equilibrium).

1.3.4 The Role of the Discount Factor

This subsection analyzes how the equilibrium payoffs of players change as the discount factor changes.

I first consider the case where $x^* > x^N$, where $(x^N, f(x^N))$ is the Nash bargaining solution outcome. The following result characterizes the SPE of the game for δ close to 1.

tinuous, strongly monotone, and quasiconcave inside the bargaining set. It is unnecessary because players make offers on the Pareto frontier in equilibrium. Therefore, only the arbitration rule defined on the Pareto frontier matters for the equilibrium outcome.

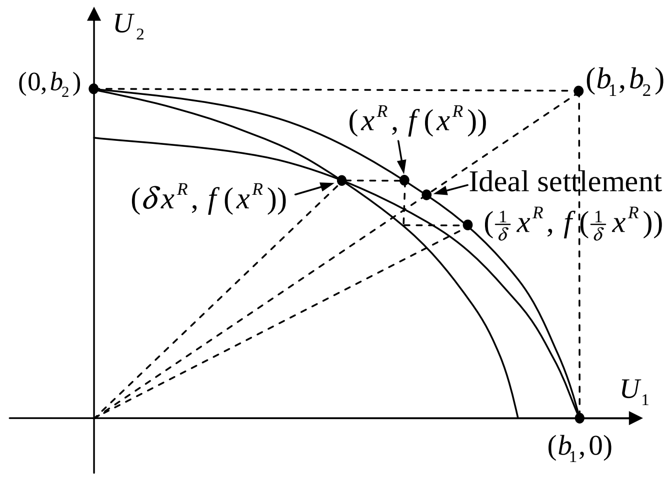

Figure 7: Sufficient condition for the Rubinstein equilibrium outcome in the KS final-offer arbitration game.

Theorem 5. *Assume that* $x^* > x^N$*. Let* $\overline{\delta}$ *be the unique* $\delta \in (0,1)$ *that satisfies* $\hat{x}(\delta) = \frac{1}{\delta}x^R(\delta)$*. Let* $\overline{\overline{\delta}}$ *be the largest* $\delta \in [0, \overline{\delta}]$ *that satisfies* $\hat{x}(\delta) = \delta x^R(\delta)$*. We have (i) if* $\overline{\delta} < \delta < 1$*, then the unique SPE of the alternating-offer arbitration game is a type-II arbitration-driven equilibrium, in which the agreement is delayed, and (ii) if* $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$*, then the unique SPE of the alternating-offer arbitration game is the Rubinstein equilibrium.*

Sketch of proof: The threshold discount factor $\overline{\delta}$ exists and is unique due to the following facts (see Figure 8): (i) $\hat{x}(\delta)$ is increasing in $\delta \in (0,1)$; (ii) $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0,1)$;[14] (iii) as δ approaches 1, $\hat{x}(\delta)$ approaches x^* and $\frac{1}{\delta}x^R(\delta)$ approaches x^N where $x^N < x^*$; and (iv) as δ approaches 0, $\frac{1}{\delta}x^R(\delta)$ goes to infinity.

Threshold $\overline{\overline{\delta}}$ is also well-defined because there is at least one point ($\delta = 0$) at which $\hat{x}(\delta) = \delta x^R(\delta)$.[15] In addition, since the curve $\delta x^R(\delta)$ is below the curve $\frac{1}{\delta}x^R(\delta)$ for

[14]See Appendix 3 for the proof.

[15]Figure 8 illustrates the case where the curve $\hat{x}(\delta)$ intersects with the curve $\delta x^R(\delta)$ exactly

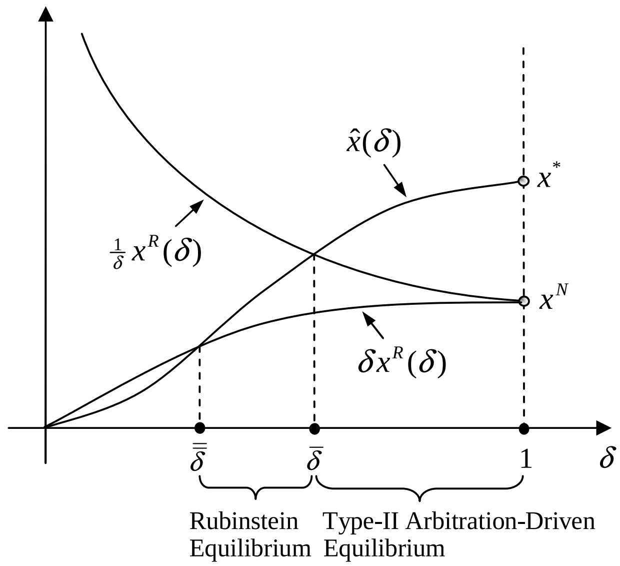

Figure 8: Determination of $\overline{\delta}$ and $\overline{\overline{\delta}}$.

all $\delta \in [0, 1)$, it must be true that as δ decreases from 1 to 0, the curve $\hat{x}(\delta)$ first intersects with the curve $\dfrac{1}{\delta}x^R(\delta)$, and then intersects with the curve $\delta x^R(\delta)$. So, we must have $\overline{\overline{\delta}} < \overline{\delta}$.

Theorem 5 then follows from the above analysis and Theorem 1. $\square$

Figure 9 illustrates the equilibrium payoffs received by players. Figure 9 reveals that the equilibrium payoff of Player 1 is $x^R(\delta)$ for $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$ and is $\delta\hat{x}(\delta)$ for $\overline{\delta} < \delta < 1$. Note that when $\delta = \overline{\delta}$, Player 1's payoff obtained from the Rubinstein equilibrium is the same as that obtained from the type-II arbitration-driven equilibrium, i.e., $x^R(\overline{\delta}) = \overline{\delta}\hat{x}(\overline{\delta})$.

As δ increases, the equilibrium payoff of Player 1 strictly decreases. This implies

once, besides at $\delta = 0$. Notice that depending on the shape of the curve $\hat{x}(\delta)$, the curve $\hat{x}(\delta)$ can intersect with the curve $\delta x^R(\delta)$ multiple times besides at $\delta = 0$. Thus, we may have either the Rubinstein equilibrium or type-I arbitration-driven equilibrium for a given $\delta \leq \overline{\delta}$, depending on whether $\hat{x}(\delta) \geq \delta x^R(\delta)$ or $\hat{x}(\delta) < \delta x^R(\delta)$. However, for $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$, we must have $\hat{x}(\delta) \geq \delta x^R(\delta)$.

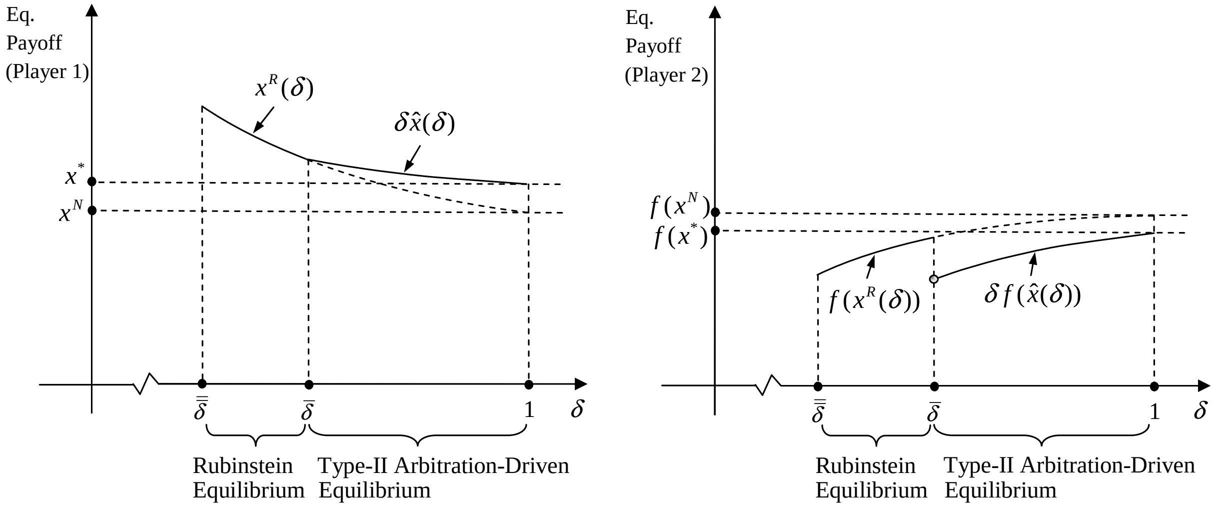

Figure 9: Players' equilibrium payoffs as functions of the discount factor.

that Player 1's payoff is smaller when both players are more patient.

In contrast, the equilibrium payoff of Player 2 is $f(x^R(\delta))$ for $\bar{\bar{\delta}} \leq \delta \leq \bar{\delta}$ and is $\delta f(\hat{x}(\delta))$ for $\bar{\delta} < \delta < 1$. Note that $f(x^R(\bar{\delta})) > \bar{\delta} f(\hat{x}(\bar{\delta}))$.

Player 1's equilibrium payoff is continuous in δ for any $\delta \in [\bar{\bar{\delta}}, 1)$. However, Player 2's equilibrium payoff is discontinuous at $\delta = \bar{\delta}$. At $\delta = \bar{\delta}$, the equilibrium of the game switches from the Rubinstein equilibrium to a type-II arbitration-driven equilibrium. The total equilibrium payoff of the two players shrinks at $\bar{\delta}$. This is because the Rubinstein equilibrium is an equilibrium with immediate agreement, whereas type-II arbitration-driven equilibrium is an equilibrium with delayed agreement. Player 1, as the player who first makes an offer, is "immune" to the switch between equilibria. However, Player 2 is vulnerable and is subject to a strict payoff loss at $\delta = \bar{\delta}$.

Manzini and Mariotti (2001) also obtain a discontinuity result for the equilibrium payoff. In their game, the equilibrium payoffs of both players are "semidiscontinuous" within a range of discount factors. However, the mechanisms behind the appearance of discontinuity are very different. In their game, the discontinuity appears due to a multiplicity of equilibria. In my game, the discontinuity appears as a result of delay in bargaining.

Finally, when $x^* \leq x^N$, it must be true that $\hat{x}(\delta) < \dfrac{1}{\delta}x^R(\delta)$ for any $\delta \in (0,1)$. Thus, for any $\delta \in (0,1)$, the unique SPE of the arbitration game is either a type-I arbitration-driven equilibrium or the Rubinstein equilibrium. In both types of equilibrium, the agreement is reached immediately. The equilibrium payoffs of both players are thus continuous in the discount factor.

1.4 Split-the-Difference Arbitration Game

This section studies the arbitration game that uses the split-the-difference arbitration rule.[16] For simplicity, I assume that players can only make offers on the Pareto frontier.

I first consider the case where the Pareto frontier is linear. I show in Theorem 6 that three types of equilibria appear as the discount factor is varied from 0 to 1. In Theorem 6, an equilibrium with immediate agreement is one in which Player 1 makes an offer that Player 2 accepts immediately. An equilibrium with delayed agreement is one in which Player 1 makes an offer that Player 2 rejects; and at the next stage, Player 2 makes an counteroffer that Player 1 accepts. An equilibrium

[16] In general, the split-the-difference arbitration rule results in an *inefficient* arbitration outcome (i.e., the arbitration outcome lies inside the bargaining set). Rong (2011) proposes a modified version of the split-the-difference rule, called the *symmetric arbitration rule*, which not only "splits the difference" between offers, but also results in an efficient arbitration outcome. Rong (2011) finds that the unique equilibrium outcome of the arbitration game using the symmetric arbitration rule coincides with the Kalai-Smorodinsky solution outcome as long as both players are sufficiently patient.

with no agreement is one in which all offers are rejected and the final outcome splits the difference between offers.

Theorem 6. *Suppose that the bargaining set S has a linear Pareto frontier with $b_1 = b_2 = 1$, i.e., $f(x) = 1 - x$ for $x \in [0, 1]$. Then the equilibrium of the split-the-difference arbitration game as a function of δ is described by Table 1.*

δ	**Equilibrium Type**	**Equilibrium Initial Offer (x_1)**	**Equilibrium Counteroffer (x_2)**
$0 < \delta \leq 0.752$	Immediate agreement	$(2-\delta)/(2+\delta)$	NA
$0.752 < \delta \leq 0.763$	Delayed agreement	1	$\delta/(2-\delta)$
$0.763 < \delta < 0.781$	Delayed agreement	$2(2-\delta)(1-\delta)/\delta^2$	$2(1-\delta)/\delta$
$0.781 \leq \delta \leq 0.868$	Immediate agreement	$(2-\delta)/(2+\delta)$	NA
$0.868 < \delta \leq 0.874$	Immediate agreement	$(2-2\delta^2)/(2-\delta^2)$	NA
$0.874 < \delta < 1$	No agreement	1	0

Table 1: SPE of the game in the split-the-difference arbitration game (linear Pareto frontier).

Proof: See Appendix 3. □

Based on the players' equilibrium strategies listed in Table 1, I depict the equilibrium payoff received by Player 1 in Figure 10. There are two interesting results in Figure 10. First, for some ranges of discount factors, as the discount factor increases, the equilibrium payoff of Player 1 increases.[17] This happens when the equilibrium is either an equilibrium with delayed agreement ($0.752 < \delta \leq 0.763$), or an equilibrium with no agreement ($0.874 < \delta < 1$). In those two types of equilibria, no agreement is reached at Stage 1. Thus, in those two equilibria, after Player 2 rejects Player 1's

[17]Notice that Player 1's payoff obtained from the Rubinstein equilibrium is always strictly decreasing in δ (see Appendix 3).

offer, the game moves to Stage 2 and Player 2 becomes the proposer. The switch of the proposer role between the two players complicates the relationship between the discount factor and the initial proposer's equilibrium payoff and makes it possible for Player 1 to increase his equilibrium payoff as the discount factor rises. In contrast, in the standard alternating-offer model that features immediate agreement (e.g., Rubinstein 1982), as players become more patient, the payoff obtained by Player 1 decreases.

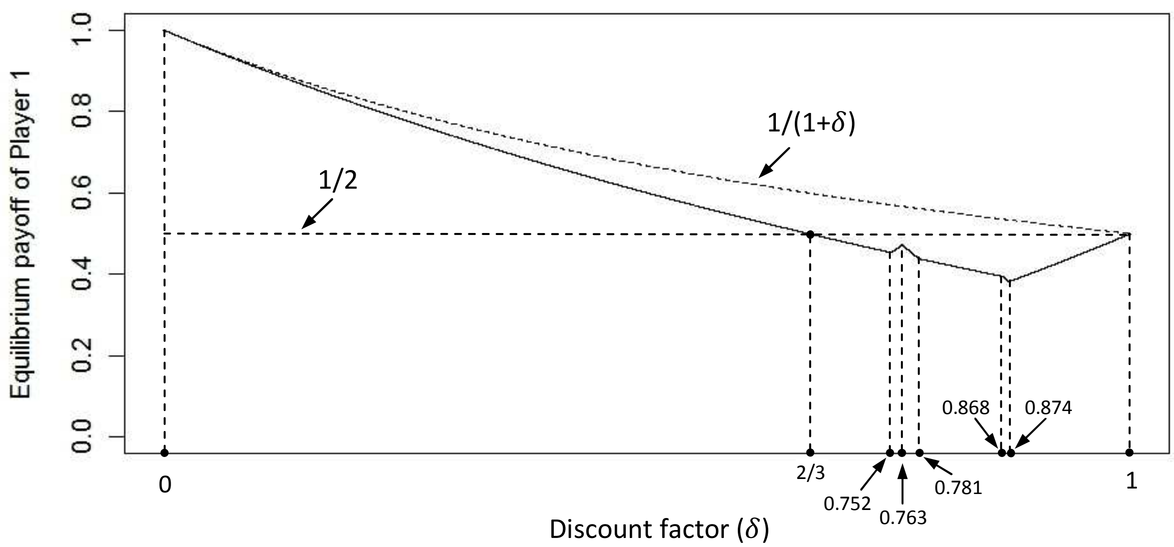

Figure 10: Equilibrium payoff to Player 1 as a function of the discount factor.

Second, for any $\delta > 2/3$, the equilibrium payoff of Player 1 is strictly less than 1/2, the fair division payoff and the Nash bargaining solution payoff of the game. This occurs even when there is an immediate agreement in the game ($2/3 < \delta \leq 0.752$ and $0.781 \leq \delta \leq 0.874$). This means that with split-the-difference arbitration, when players become patient, Player 1's bargaining power becomes "less" than Player 2's bargaining power, even though Player 1 makes the first offer. The reason is as follows.

As in the standard alternating-offer model, Player 1 (as the player who makes the first offer) has the first-mover advantage because the players are impatient. On the other hand, Player 2 has the second-mover advantage of being "closer" to the arbitration stage, and thus Player 2 can more credibly threaten to make the extreme demand in order to obtain a favorable outcome if the arbitration stage is reached. When the players become patient, Player 1's first-mover advantage decreases, while Player 2's second-mover advantage increases. It is thus not surprising that when players become sufficiently patient, Player 1's bargaining power becomes "less" than Player 2's.

Figure 10 also shows that Player 1's equilibrium payoff is consistently less than the Rubinstein equilibrium payoff.

Table 1 shows that the unique SPE of the game depends on the discount factor in a complex manner. However, the equilibrium features immediate agreement when the discount factor is sufficiently small, and it is an equilibrium with no agreement when the discount factor is sufficiently large. This turns out to be a general property that holds even when the Pareto frontier is not linear. This result is summarized in Theorem 18.

For any $\delta \in (0,1)$, define $x_1^*(\delta)$ as the unique $x_1 \in (0, b_1)$ that satisfies $f(x_1) = \delta f(\dfrac{\delta}{2-\delta} x_1)$. We have:

Theorem 7. *In the split-the-difference arbitration game, there exists two thresholds $\delta_1^*, \delta_2^* \in (0,1)$ with $\delta_1^* > \delta_2^*$, such that (i) when $\delta \in (\delta_1^*, 1)$, the only SPE of the game is that at Stage 1, Player 1 offers $(b_1, 0)$, which Player 2 rejects; and at Stage 2, Player 2 offers $(0, b_2)$, which Player 1 rejects, and (ii) when $\delta \in (0, \delta_2^*)$, the only*

SPE of the game is that at Stage 1, Player 1 offers $(x_1^*(\delta), f(x_1^*(\delta)))$*, which Player 2 accepts immediately.*

Proof: See Appendix 2. $\square$

Theorem 18 states that if δ is sufficiently large, then the unique SPE of the game is such that both players make extreme offers. This occurs because, as δ becomes sufficiently large, the time cost becomes so low that each player would rather choose to make the extreme offer when it is his turn to make the offer. In particular, by making the extreme offer, a player can guarantee himself a payoff of approximately half of the maximum payoff that he can obtain from the bargaining set. However, by making an offer that will be accepted by the other player, a player can only obtain a small payoff because the other player can always threaten to reject the offer, make an extreme counteroffer and move the game to arbitration if the offer is not good enough (the threat is credible since players are very patient). Theorem 18 also states that if δ is sufficiently small, then the unique SPE of the game is an equilibrium with immediate agreement.

The arbitration stage is never reached in equilibrium in the final-offer arbitration game, but it might be reached in equilibrium in the split-the-difference arbitration game. The reason is that, unlike final-offer arbitration, split-the-difference arbitration chooses a compromise between final offers as the arbitration outcome. This feature of split-the-difference arbitration encourages players to make extreme offers before arbitration (and move the game to the arbitration stage).[18] The incentives

[18]This is also known as the *chilling effect* of conventional arbitration, which holds for any arbitration mechanism that allows for compromise between offers (Feuille 1975; Deck and Farmer 2007).

for players to make extreme offers become stronger as they grow more patient.

1.5 Conclusion

This paper studies a finite-horizon alternating-offer model that involves arbitration. I find that when final-offer arbitration is used, there exists a wide range of arbitrator preferences (i.e., the set of balanced final-offer arbitration rules), under which the unique equilibrium outcome of the arbitration game is unaffected by the specific details of the arbitrator's preference. Within this range, the equilibrium outcome coincides with the Rubinstein equilibrium outcome. Outside this range, delay in equilibrium might arise. If, instead, the arbitration rule splits the difference between offers, then the unique equilibrium of the game depends on the discount factor. In particular, when the discount factor is sufficiently small, the unique SPE must be an equilibrium with immediate agreement; and when the discount factor is sufficiently large, the unique SPE must be an equilibrium with no agreement (and the arbitration stage will be reached).

Both the balanced final-offer arbitration rules and the split-the-difference arbitration rule might be regarded as fair arbitration rules. However, they have very different implications for players' equilibrium payoffs. In particular, if a balanced final-offer arbitration rule is used, players always obtain Rubinstein equilibrium payoffs. If, instead, the split-the-difference arbitration rule is used, then Player 1's payoff in the arbitration game is always less than the Rubinstein equilibrium payoff (when the Pareto frontier is linear). These results reflect that Player 2 can credibly threaten to make the extreme offer in the split-the-difference arbitration game, but not in the balanced final-offer arbitration game.[19] As a result, Player 1's bargaining power in the split-the-difference arbitration game is "less" than that in the final-offer arbitration game and Player 1 obtains a smaller payoff in the split-the-difference arbitration game.

A crucial feature of the arbitration games considered in this paper is that the arbitration outcome depends on players' offers. This dependency distinguishes my model from the outside option literature[20] in terms of equilibrium strategies and equilibrium outcomes. The differences include (i) in my model, Player 2's optimal counteroffer depends on Player 1's initial offer, so that Player 1 can control Player 2's counteroffer by varying his own offer; (ii) the condition that yields the Rubinstein equilibrium outcome in my model is different from that obtained by Manzini and Mariotti (2001), and (iii) delayed agreements can occur in my model even though there is always a unique equilibrium.

1.6 Chapter 1 Appendix

1.6.1 Appendix 1: Proof of Theorem 1

I use the following four lemmas (Lemma 8, Lemma 9, Lemma 10 and Lemma 11) to prove Theorem 1.

[19]By making the extreme offer in the split-the-difference arbitration game, a player guarantees himself a payoff of at least half of the maximum payoff he can obtain from the bargaining set, while making the extreme offer in the arbitration game with a balanced final-offer arbitration rule will usually lead the arbitrator to choose the other player's offer.

[20]See the joint outside option model considered by Manzini and Mariotti (2001, 2004) and the unilateral outside option model considered by Binmore et al. (1989), Ponsatí and Sákovics (1998) and Shaked (1994). In the outside option literature, the outside options of players are exogenously given.

Lemma 8. *If the final offer arbitration rule h is regular (i.e., $h \in \Sigma$), then*

(i) for any $x_2 \in [g(x_1), x_1]$, we have $h((x_1, f(x_1)), (x_2, f(x_2))) = (x_2, f(x_2))$;

(ii) for any $x_2 \in [0, g(x_1))$, we have $h((x_1, f(x_1)), (x_2, f(x_2))) = (x_1, f(x_1))$.

Sketch of Proof: The lemma follows from the definition of $g(x_1)$, and the strong monotonicity of the arbitrator's utility function. Notice that since $x_1 \in V_x(x_1, f(x_1))$, it must be true that $g(x_1) \leq x_1$ and the interval $[g(x_1), x_1]$ is well-defined. □

According to Lemma 8, for any given offer made by Player 1 $(x_1, f(x_1))$, the best counteroffer that Player 2 could make on the Pareto frontier and the arbitrator would choose is $(g(x_1), f(g(x_1)))$.

The following lemma characterizes Player 2's best counteroffer at Stage 2, generalizing a result in Yildiz (2011).

Lemma 9. *In the arbitration game where $h \in \Sigma$, if Player 1 offers $(x_1, y_1) \in PF$ at Stage 1 and Player 2 rejects it, then at Stage 2, in any equilibrium subgame, Player 2 makes the offer $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$ and Player 1 accepts it.*

Proof: I first show that it is never optimal for Player 2 to make an offer that is strictly inside the bargaining set (i.e., not on the Pareto frontier). To do this, I establish the following two facts. First, it is never optimal for Player 2 to make an offer that is rejected by Player 1. Second, for any Player 2's offer $(x_2, y_2) \notin PF$ that Player 1 would accept, the offer (x_2, y_2) must be strictly dominated by the offer $(x_2, f(x_2))$, which is on the Pareto frontier.

To establish the first point, notice that if Player 2's offer is rejected by Player 1, then it must be true that Player 1's offer is the arbitration outcome. Thus, Player 2 would be strictly better off if he offers $(x_2, y_2) = (x_1, y_1)$, which will be accepted by Player 1 immediately.

To establish the second point, suppose that $(x_2, y_2) \notin PF$ would be accepted by Player 1. Then, it must be true that $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ or $x_2 \geq \delta x_1$ (or both). Since the arbitrator's utility function is strongly monotone, $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ implies that $h((x_1, y_1), (x_2, f(x_2))) = (x_2, f(x_2))$. Thus, if Player 2 offers $(x_2, f(x_2))$, it would also be accepted by Player 1. But then Player 2 obtains a strictly higher payoff by offering $(x_2, f(x_2))$.

It follows that Player 2 never makes an offer that is strictly inside the bargaining set. Without loss of generality, I now restrict Player 2's offer to be on the Pareto frontier. We have the following two possibilities.

(i) Player 2 offers $(x_2, y_2) \in PF$ with $x_2 < \min\{\delta x_1, g(x_1)\}$. If Player 1 accepts the offer, then his payoff is δx_2 (the payoff is measured at Stage 1). If Player 1 rejects the offer, the game moves to the arbitration stage. Given that $x_2 < g(x_1)$, we must have $h((x_1, y_1), (x_2, y_2)) = (x_1, y_1)$ (by Lemma 8 (ii)). Thus, Player 1's payoff is $\delta^2 x_1$. Since $x_2 < \delta x_1$, we have $\delta x_2 < \delta^2 x_1$, so Player 1 will reject Player 2's offer and Player 2's payoff is $\delta^2 y_1$.

(ii) Player 2 offers $(x_2, y_2) \in PF$ with $x_2 \geq \min\{\delta x_1, g(x_1)\}$. If Player 1 accepts, his payoff is δx_2. If Player 1 rejects the offer, the game moves to the arbitration stage. If (x_2, y_2) is the arbitrated outcome, then Player 1's payoff is $\delta^2 x_2$, which is less than δx_2. If instead, (x_1, y_1) is the arbitrated outcome, then Player 1's payoff is $\delta^2 x_1$. In this latter case, since (x_1, y_1) is the arbitrated outcome, we must have either $x_2 < g(x_1)$ or $x_2 \geq x_1$ (by Lemma 8). If $x_2 < g(x_1)$, noting that $x_2 \geq \min\{\delta x_1, g(x_1)\}$, we must have $x_2 \geq \delta x_1$, which implies that $\delta^2 x_1 \leq \delta x_2$. If $x_2 \geq x_1$, then we have $\delta^2 x_1 \leq \delta x_2$. In each of these cases, Player 1 obtains a higher payoff by accepting Player 2's offer. Thus, Player 1 will accept Player 2's offer and Player 2's payoff is δy_2.[21]

In summary, if Player 2 offers $(x_2, y_2) \in PF$ with $x_2 < \min\{\delta x_1, g(x_1)\}$, then his equilibrium payoff is $\delta^2 y_1 \leq \delta y_1 = \delta f(x_1) < \delta f(\min\{\delta x_1, g(x_1)\})$, where the last inequality follows from the fact that $\min\{\delta x_1, g(x_1)\} \leq \delta x_1 < x_1$.[22] If Player 2 offers $(x_2, y_2) \in PF$ with $x_2 \geq \min\{\delta x_1, g(x_1)\}$, then his equilibrium payoff is δy_2, which is maximized at $(x_2, y_2) = (\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$ with the corresponding payoff for Player 2 being $\delta f(\min\{\delta x_1, g(x_1)\})$. Comparing these two cases, it is obvious that Player 2's optimal counteroffer is $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$. Moreover, Player 1 will accept the offer $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$. $\square$

Three factors determine Player 2's best counteroffer at Stage 2: (i) the discount factor δ; (ii) the final-offer arbitration rule h; and (iii) Player 1's initial offer $(x_1, f(x_1))$[23].

Define $\tilde{x}(x_1) = \min\{\delta x_1, g(x_1)\}$. Define the *(optimal) counteroffer curve* (of Player 2) as the curve $x = \tilde{x}(f^{-1}(y))$ where $y \in [0, b_2]$ (see Figure 11). The counteroffer curve is the collection of points $(\tilde{x}(x_1), f(x_1))$ as x_1 varies from 0 to b_1. The counteroffer curve can be used to determine Player 2's optimal counteroffer for any given Player 1's offer, if Player 2 chooses to reject Player 1's offer. See Figure 11.

[21] Here, we used tie-breaking rule 1.

[22] Note that the last inequality is strict because $x_2 < \min\{\delta x_1, g(x_1)\}$ implies $x_1 > 0$.

[23] The dependency of Player 2's counteroffer on Player 1's initial offer is a key feature of my arbitration game. This dependency is absent in Rubinstein's infinite-horizon alternating-offer game.

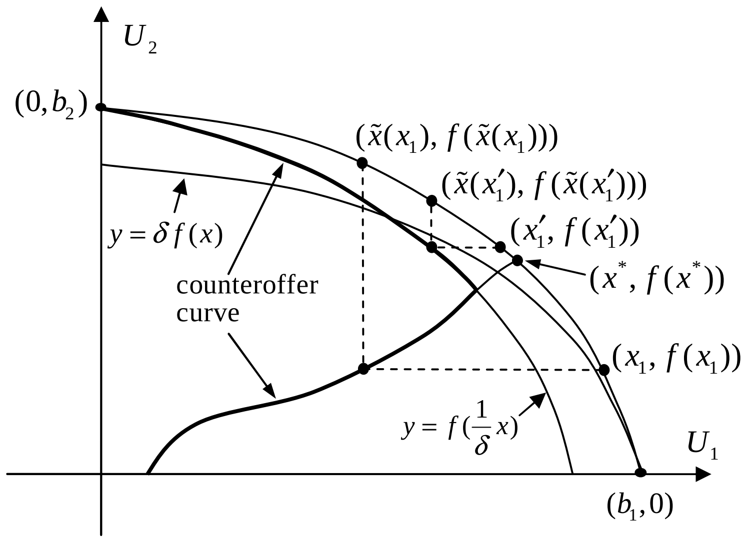

Figure 11: The counteroffer curve of Player 2.

The next lemma generalizes Lemma 9. It characterizes Player 2's optimal counteroffer when Player 1's offer (x_1, y_1) is not on the Pareto frontier. Its proof is similar to that of Lemma 9 and is omitted.

Let $g(x_1, y_1) = \min\{x_2 | x_2 \in V_x(x_1, y_1)\}$ and $\tilde{x}(x_1, y_1) = \min\{\delta x_1, g(x_1, y_1)\}$.[24] We have:

Lemma 10. *In the arbitration game with $h \in \Sigma$, if Player 1 made an offer (x_1, y_1) at Stage 1 and Player 2 rejected it, then at Stage 2, in any equilibrium subgame, Player 2 makes the offer $(\tilde{x}(x_1, y_1), f(\tilde{x}(x_1, y_1)))$ and Player 1 accepts it.*

The following lemma characterizes the necessary and sufficient conditions for Player 1's offer $(x_1, y_1) \in PF$ to be accepted by Player 2.

Lemma 11. *In the arbitration game with $h \in \Sigma$, if Player 1 made an offer $(x_1, y_1) \in PF$ at Stage 1, then in equilibrium:*

[24] Notice that $g(x_1) = g(x_1, f(x_1))$ and $\tilde{x}(x_1) = \tilde{x}(x_1, f(x_1))$.

(i) If $\delta x^R \leq \hat{x} \leq \delta b_1$, then Player 1's offer (x_1, y_1) is accepted by Player 2 if and only if $0 \leq x_1 \leq x^R$;

(ii) If $0 \leq \hat{x} < \delta x^R$, then Player 1's offer (x_1, y_1) is accepted by Player 2 if and only if $0 \leq x_1 \leq f^{-1}(\delta f(\hat{x}))$.

Proof:

(i) Suppose $\hat{x}$ is such that $\delta x^R \leq \hat{x} \leq \delta b_1$.

Refer to Figure 12. There are two cases.

(a) At Stage 1, Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$. If Player 2 accepts the offer, then his payoff is $f(x_1)$. If Player 2 rejects the offer, then at Stage 2, from Lemma 9, he offers $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$ that Player 1 accepts; Player 2's payoff is $\delta f(\tilde{x}(x_1))$. We have $\delta f(\tilde{x}(x_1)) \geq \delta f(\delta x_1) > f(x_1)$, where the first inequality follows from the fact that $\tilde{x}(x_1) = \min\{\delta x_1, g(x_1)\} \leq \delta x_1$ and the second inequality follows from the fact that $x_1 > x^R$ (see Figure 12). So, Player 2 rejects the offer $(x_1, f(x_1))$ and makes the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 accepts.

(b) At Stage 1, Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq x^R$. If Player 2 accepts the offer, then his payoff is $f(x_1')$. If Player 2 rejects the offer, then at Stage 2, from Lemma 9, he offers $(\tilde{x}(x_1'), f(\tilde{x}(x_1')))$, which Player 1 accepts. Thus, Player 2's payoff is $\delta f(\tilde{x}(x_1')) = \delta f(\delta x_1') \leq f(x_1')$ (see Figure 12). Consequently, Player 2 accepts $(x_1', f(x_1'))$.

(ii) *Suppose $\hat{x}$ is such that $0 \leq \hat{x} < \delta x^R$.*

Refer to Figure 13. There are two cases.

(a) At Stage 1, Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > f^{-1}(\delta f(\hat{x}))$. If Player 2 accepts the offer, then his payoff is $f(x_1)$. If Player 2 rejects the offer, then at Stage

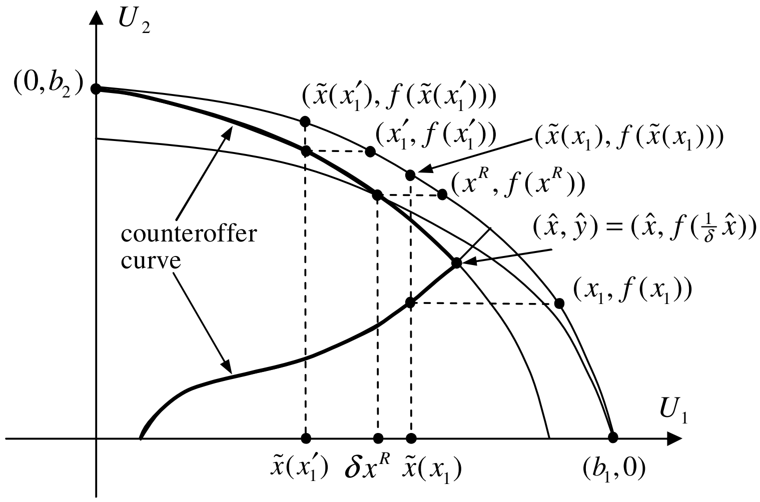

Figure 12: The case of $\delta x^R \leq \hat{x} \leq \delta b_1$.

2, from Lemma 9, he offers $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$ that Player 1 accepts. Thus, Player 2's payoff is $\delta f(\tilde{x}(x_1))$. We have $\delta f(\tilde{x}(x_1)) \geq \delta f(\hat{x}) > f(x_1)$, where the first inequality follows from the fact that $\tilde{x}(x_1) = \min\{\delta x_1, g(x_1)\} = g(x_1) \leq \hat{x}$ (see Figure 13) and the second inequality follows from the fact that $x_1 > f^{-1}(\delta f(\hat{x}))$. Therefore, Player 2 rejects the offer $(x_1, f(x_1))$ and makes the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 accepts.

(b) At Stage 1, Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq f^{-1}(\delta f(\hat{x}))$. If Player 2 accepts the offer, then his payoff is $f(x_1')$. If Player 2 rejects the offer, then at Stage 2, from Lemma 9, he offers $(\tilde{x}(x_1'), f(\tilde{x}(x_1')))$, which Player 1 accepts; Player 2's payoff is $\delta f(\tilde{x}(x_1'))$. Since $\delta f(\tilde{x}(x_1')) \leq f(x_1')$ (see Figure 13), Player 2 will accept the offer $(x_1', f(x_1'))$. $\square$

Now, we can state the proof of Theorem 1.

Proof of Theorem 1:

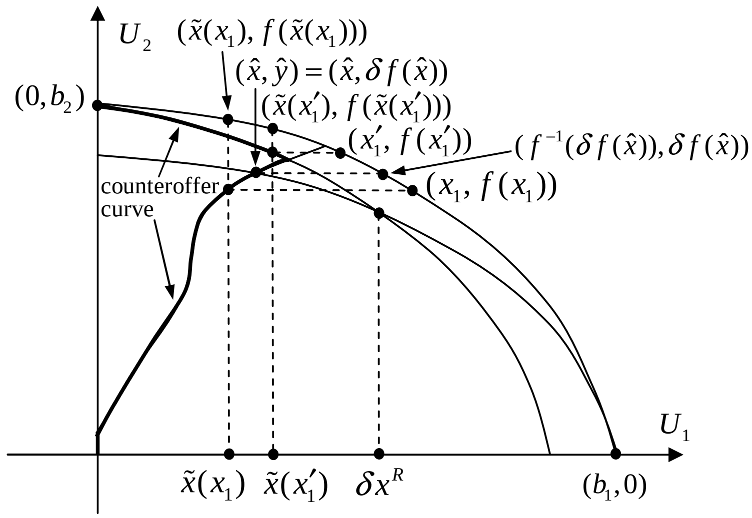

Figure 13: The case of $0 \le \hat{x} < \delta x^R$.

I first show that it is never optimal for Player 1 to make an offer that is strictly inside the bargaining set (i.e., not on the Pareto frontier). In particular, I show that if Player 1 offers $(x_1, y_1) \notin PF$, then the offer is strictly dominated by the offer $(f^{-1}(y_1), y_1)$.

Suppose (x_1, y_1) would be accepted by Player 2. Then, it must be true that $y_1 \ge \delta f(\tilde{x}(x_1, y_1))$. Since the arbitrator's utility function is strongly monotone, we have $g(f^{-1}(y_1), y_1) > g(x_1, y_1)$ and $\tilde{x}(f^{-1}(y_1), y_1) > \tilde{x}(x_1, y_1)$. So, $y_1 \ge \delta f(\tilde{x}(x_1, y_1))$ implies that $y_1 \ge \delta f(\tilde{x}(f^{-1}(y_1), y_1))$. Thus, if Player 1 makes the offer $(f^{-1}(y_1), y_1)$, then it will also be accepted by Player 2. Player 1 thus obtains a strictly higher payoff by offering $(f^{-1}(y_1), y_1)$.

Suppose (x_1, y_1) would be rejected by Player 2. Then, Player 1 must obtain a payoff of $\delta\tilde{x}(x_1, y_1)$. If Player 1 makes the offer $(f^{-1}(y_1), y_1)$, then it might be accepted by Player 2, or rejected by Player 2. If Player 2 accepts the offer, then

Player 1 obtains a payoff of $f^{-1}(y_1) > x_1 \geq \delta\tilde{x}(x_1, y_1)$, where the last inequality follows from the fact that $\delta\tilde{x}(x_1, y_1) \leq \delta(\delta x_1) \leq x_1$. If Player 2 rejects the offer, then Player 1 obtains a payoff of $\delta\tilde{x}(f^{-1}(y_1), y_1)$, which is strictly greater than $\delta\tilde{x}(x_1, y_1)$.

In all the cases analyzed above, Player 2 is strictly better off by making the offer $(f^{-1}(y_1), y_1)$. I thus showed that Player 1 never makes an offer that is strictly inside the bargaining set. In the remainder of the proof, I assume that Player 1 can only make an offer on the Pareto frontier. There are three cases.

(i) Suppose $\hat{x}$ is such that $\delta x^R \leq \hat{x} \leq \frac{1}{\delta}x^R$.

If Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$, then Player 2 will reject it and make the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 will accept (by Lemma 11 (i) and Lemma 9). Player 1's payoff is thus $\delta\tilde{x}(x_1)$, which is maximized at $x_1 = \frac{1}{\delta}\hat{x}$ with corresponding payoff $\delta\tilde{x}(\frac{1}{\delta}\hat{x}) = \delta\hat{x}$.[25] If Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq x^R$, then Player 2 will accept it (by Lemma 11 (i)). Player 1's payoff is x_1', which is maximized at $x_1' = x^R$ with corresponding payoff x^R. Since $\delta\hat{x} \leq x^R$, Player 1's optimal strategy is to offer $(x^R, f(x^R))$ at Stage 1.[26] Player 2 accepts $(x^R, f(x^R))$ immediately.

(ii) Suppose $\hat{x}$ is such that $\frac{1}{\delta}x^R < \hat{x} \leq \delta b_1$.

If Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$, then Player 2 will reject the offer $(x_1, f(x_1))$ and make the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 will accept (by Lemma 11 (i) and Lemma 9). Player 1's payoff is thus $\delta\tilde{x}(x_1)$, which is

[25] This is true except when $\hat{x} = \delta x^R$, where the maximum of $\delta\tilde{x}(x_1)$ does not exist. However, in this case, one can show that for Player 1, any offer $(x_1, f(x_1))$ with $x_1 > x^R$ must be strictly dominated by the offer $(x^R, f(x^R))$, and thus Player 1 never makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$.

[26] If $\delta\hat{x} = x^R$, Player 1 is indifferent between offering $(x^R, f(x^R))$ and offering $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$. Using tie-breaking rule 2, Player 1 must offer $(x^R, f(x^R))$.

maximized at $x_1 = \frac{1}{\delta}\hat{x}$ with corresponding payoff $\delta\tilde{x}(\frac{1}{\delta}\hat{x}) = \delta\hat{x}$. If Player 1 offers $(x_1', f(x_1'))$ with $0 \leq x_1' \leq x^R$, then Player 2 will accept it (by Lemma 11 (i)). Player 1's payoff is x_1', which is maximized at $x_1' = x^R$ with corresponding payoff x^R. Since $\delta\hat{x} > x^R$, Player 1's optimal strategy is to offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$. Player 2 will reject the offer and make the counteroffer $(\tilde{x}(\frac{1}{\delta}\hat{x}), f(\tilde{x}(\frac{1}{\delta}\hat{x}))) = (\hat{x}, f(\hat{x}))$, which Player 1 will accept.

(iii) Suppose $\hat{x}$ is such that $0 \leq \hat{x} < \delta x^R$.

If Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > f^{-1}(\delta f(\hat{x}))$, then Player 2 will reject the offer $(x_1, f(x_1))$ and make the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 will accept (by Lemma 11 (ii) and Lemma 9). Player 1's payoff is thus $\delta\tilde{x}(x_1)$, which is at most $\delta\hat{x}$ (see Figure 13). If Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq f^{-1}(\delta f(\hat{x}))$, then Player 2 will accept the offer (by Lemma 11 (ii)). Player 1's payoff is x_1', which is maximized at $x_1' = f^{-1}(\delta f(\hat{x}))$ with corresponding payoff $f^{-1}(\delta f(\hat{x}))$. Since $f^{-1}(\delta f(\hat{x})) > \delta\hat{x}$ (see Figure 13), Player 1's optimal strategy is to offer $(x_1, f(x_1))$ where $x_1 = f^{-1}(\delta f(\hat{x}))$. Player 2 will accept the offer immediately.
□

1.6.2 Appendix 2: Proof of Theorem 18

In this appendix, I first propose and prove three lemmas (Lemma 15, Lemma 17, and Lemma 16). I then give the formal proof of Theorem 18 using the three lemmas.

I begin with a definition.

Definition 12. *For any given* $(x_1, y_1) \in PF$*, define* $\hat{x}_2(x_1, y_1) = \dfrac{\delta}{2-\delta}x_1$.

The following lemma characterizes Player 2's optimal action at Stage 2 if he rejects Player 1's offer at Stage 1.

Lemma 13. *In the equilibrium of the split-the-difference arbitration game, if Player 2 rejects Player 1's offer* $(x_1, y_1) \in PF$ *at Stage 1, then at Stage 2, Player 2 must either offer* $(0, b_2)$*, which Player 1 will reject, or offer* $(\hat{x}_2(x_1, y_1), f(\hat{x}_2(x_1, y_1)))$*, which Player 1 will accept.*

Proof: Using the definition of $\hat{x}_2(x_1, y_1)$, we have $\delta \hat{x}_2(x_1, y_1) \geq \delta^2 \dfrac{\hat{x}_2(x_1, y_1) + x_1}{2}$ if and only if $x_2 \geq \hat{x}_2(x_1, y_1)$. That is, Player 1 accepts Player 2's offer (x_2, y_2) if and only if $x_2 \geq \hat{x}_2(x_1, y_1)$. Thus, if Player 2 wants to make an offer that Player 1 will accept, his best option is to offer $(\hat{x}_2(x_1, y_1), f(\hat{x}_2(x_1, y_1)))$. In addition, if Player 2 wants to make an offer (x_2, y_2) that Player 1 will reject, his best option is to make the extreme offer (i.e., $(0, b_2)$), because the arbitrated payoff $\dfrac{y_1 + y_2}{2}$ received by Player 2 is strictly increasing in y_2. □

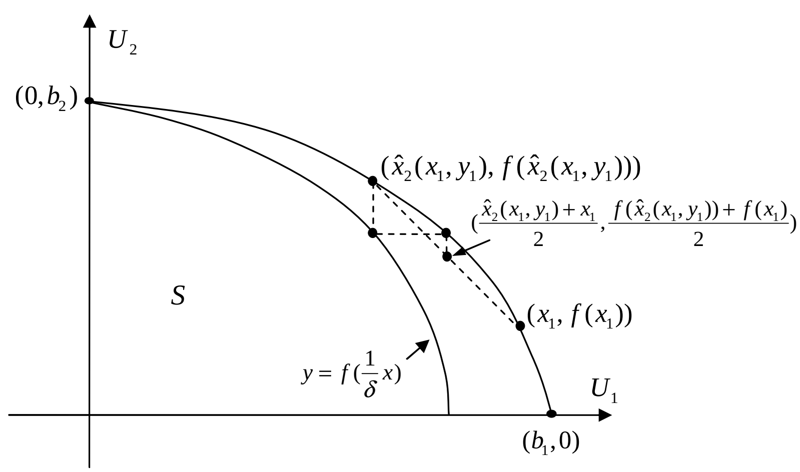

Figure 14: Definition of $\hat{x}_2(x_1, y_1)$.

For any offer $(x_1, y_1) \in PF$ made by Player 1, Player 2 can always make two threats. One threat is the counteroffer $(\hat{x}_2(x_1, y_1), f(\hat{x}_2(x_1, y_1)))$, which Player 1 will accept. The other threat is the extreme offer $(b_2, 0)$, which Player 1 will reject.[27]

Using tie-breaking rules 1 and 2, one can show that the SPE of the alternating-offer arbitration game is unique. In addition, using Lemma 15, one can show that there are at most three types of SPE in the game. These results are summarized in the following lemma.

Lemma 14. *In the split-the-difference arbitration game, the unique SPE must take one of the following forms:*

(i) (immediate agreement) at Stage 1, Player 1 offers $(x_1, y_1) \in PF$*, which Player 2 accepts;*

(ii) (delayed agreement) at Stage 1, Player 1 offers $(x_1, y_1) \in PF$*, which Player 2 rejects; and at Stage 2, Player 2 offers* $(\hat{x}_2(x_1, y_1), f(\hat{x}_2(x_1, y_1)))$*, which Player 1 accepts;*

(iii) (no agreement) at Stage 1, Player 1 offers $(x_1, y_1) \in PF$*, which Player 2 rejects; and at Stage 2, Player 2 offers* $(0, b_2)$*, which Player 1 rejects.*

The following lemma is essential for the proof of Theorem 18, and can be used to simplify the analysis of the SPE of the whole game. Notice that Lemma 15 implies

[27] The multiple threats facing Player 1 is a key feature of the split-the-difference arbitration game. Actually, this feature holds as soon as the arbitration rule is such that it chooses a compromise between offers as the arbitration outcome. This feature is absent in the game where the arbitration rule is final-offer arbitration. In the case of final-offer arbitration, at Stage 2, it is never optimal for Player 2 to make an offer that is rejected by Player 1. Due to the multiple threats, Player 1's bargaining power in the game will be significantly undermined. This point is also illustrated in the case where the Pareto frontier is linear (see the analysis of Figure 10).

that in equilibrium, if Player 1 offers $(x_1, y_1) \in PF$ at Stage 1, then Player 2 chooses among one of the following three options:

(A) accept the offer (x_1, y_1);

(B) reject (x_1, y_1), and at Stage 2, offer $(\hat{x}_2(x_1, y_1), f(\hat{x}_2(x_1, y_1)))$, which Player 1 will accept;

(C) reject (x_1, y_1), and at Stage 2, offer $(0, b_2)$, which Player 1 will reject.

Lemma 15. *In the split-the-difference arbitration game, in any equilibrium, if Player 1 offers $(x_1, f(x_1)) \neq (b_1, 0)$ at Stage 1, then Player 2 must be indifferent between his best two options among A, B and C, i.e., either $A \sim_2 B \succeq_2 C$, or $A \sim_2 C \succeq_2 B$, or $B \sim_2 C \succeq_2 A$.*

Proof: For Player 1, the payoffs of the three options (A, B and C) are x_1, $\delta\frac{\delta}{2-\delta}x_1$ and $\delta^2\frac{x_1}{2}$ respectively. For Player 2, the payoff of the three options are $f(x_1)$, $\delta f(\frac{\delta}{2-\delta}x_1)$ and $\delta^2\frac{f(x_1)+b_2}{2}$.

Note the following two facts. First, Player 1's payoff from each option is strictly increasing in x_1. Second, Player 2's payoff from each option is continuous in x_1. Using these two facts, it follows that if in equilibrium, Player 1 makes an offer (x_1, y_1) with $x_1 < b_1$ and Player 2 has a strict preference over his best two options, then Player 1's offer (x_1, y_1) cannot be an equilibrium offer. This is because Player 1 can obtain a strictly higher payoff by making a slightly more extreme offer $(x_1 + \epsilon, f(x_1 + \epsilon))$ where $\epsilon > 0$ is small enough such that Player 2's preferences over the three options are the same as before.

Thus, I have proved that, if at Stage 1 Player 1 offers $(x_1, y_1) \neq (b_1, 0)$, then Player 2 must be indifferent between the best two options of A, B and C. $\square$

Now, we can state the proof of Theorem 18.

Proof of Theorem 18:

(i) I show that there exists a $\delta_1^* \in (0,1)$ such that when $\delta \in (\delta_1^*, 1)$, the only SPE of the alternating-offer arbitration game is that at Stage 1, Player 1 offers $(b_1, 0)$, which Player 2 rejects; and at Stage 2, Player 2 offers $(0, b_2)$, which Player 1 rejects. The threshold δ_1^* is determined below.

Suppose $\delta > \delta_1^*$.

Step 1: I first show that at Stage 1, Player 1 offers $(b_1, 0)$.

If Player 1 offers $(b_1, 0)$, then he can guarantee himself a payoff of $\delta^2 \frac{b_1}{2}$. If Player 1 offers $(x_1, y_1) \in PF$ with $x_1 < b_1$, then by Lemma 16, either $A \sim_2 B \succeq_2 C$, or $A \sim_2 C \succ_2 B$, or $B \sim_2 C \succ_2 A$; I discuss these three cases below.

If one of the first two cases hold, then Player 2 accepts (x_1, y_1) and Player 1's payoff is x_1.[28] In addition, we have $A \succeq_2 C$, i.e., $f(x_1) \geq \delta^2 \frac{f(x_1) + b_2}{2}$. This implies that $f(x_1) \geq \frac{\delta^2 b_2}{2 - \delta^2}$, i.e., $x_1 \leq f^{-1}(\frac{\delta^2 b_2}{2 - \delta^2})$. Define $\delta_{11}^* \in (0,1)$ as the unique $\delta \in (0,1)$ such that $f^{-1}(\frac{\delta^2 b_2}{2 - \delta^2}) = \delta^2 \frac{b_1}{2}$. Then, for $\delta > \delta_{11}^*$, we have $x_1 \leq f^{-1}(\frac{\delta^2 b_2}{2 - \delta^2}) < \delta^2 \frac{b_1}{2}$, i.e., for Player 1, the offer (x_1, y_1) is strictly dominated by the offer $(b_1, 0)$.

If the third case holds, then (x_1, y_1) will be rejected by Player 2 and Player 2 will

[28]Using tie-breaking rule 1.

offer $(\hat{x}_2(x_1, y_1), f(\hat{x}_2(x_1, y_1)))$ at Stage 2.[29] Player 1's payoff is $\delta\hat{x}_2(x_1, y_1) = \frac{\delta^2}{2-\delta}x_1$. Since $B \sim_2 C$, we have $\delta f(\frac{\delta}{2-\delta}x_1) = \delta^2 \frac{f(x_1)+b_2}{2}$. That is, we have $x_1 = x_1^{**}(\delta)$, where $x_1^{**}(\delta)$ is the unique $x_1 \in (0, b_1)$ that satisfies $\delta f(\frac{\delta}{2-\delta}x_1) = \delta^2 \frac{f(x_1)+b_2}{2}$. Since $x_1^{**}(\delta) \to 0$ as $\delta \to 1$, there exists $\delta_{12}^* \in (0, 1)$ such that if $\delta > \delta_{12}^*$, then $\frac{\delta^2}{2-\delta}x_1^{**}(\delta) < \delta^2 \frac{b_1}{2}$, i.e., for Player 1, the offer (x_1, y_1) is strictly dominated by the offer $(b_1, 0)$. Thus, we have proved that, for $\delta > \max\{\delta_{11}^*, \delta_{12}^*\}$, Player 1 must make the offer $(b_1, 0)$.

Step 2: I now show that at Stage 2, Player 2 offers $(0, b_2)$.

First, notice that Player 2 must reject Player 1's offer $(b_1, 0)$. At Stage 2, according to Lemma 15, Player 2 either offers $(\hat{x}_2(b_1, 0), f(\hat{x}_2(b_1, 0)))$, which Player 1 accepts, or Player 2 offers $(0, b_2)$, which Player 1 rejects. Note that $\hat{x}_2(b_1, 0) = \frac{\delta}{2-\delta}b_1$. So, for Player 2, the payoffs of the above two options are $\delta f(\frac{\delta}{2-\delta}b_1)$ and $\delta^2 \frac{b_2}{2}$ respectively. Define δ_{13}^* as the unique $\delta \in (0, 1)$ such that $f(\frac{\delta}{2-\delta}b_1) = \delta \frac{b_2}{2}$. Then, for $\delta > \delta_{13}^*$, we have $\delta f(\frac{\delta}{2-\delta}b_1) < \delta^2 \frac{b_2}{2}$. That is, for all $\delta > \delta_{13}^*$, Player 2 offers $(0, b_2)$ at Stage 2.

Finally, define $\delta_1^* = \max\{\delta_{11}^*, \delta_{12}^*, \delta_{13}^*\}$ and this completes the proof.

(ii) Define $\delta_2^* = \min\{\delta_{13}^*, \delta_{21}^*, \delta_{22}^*\}$. δ_{13}^* is defined above. $\delta_{21}^* = \min_{0 \le x_1 \le b_1} \delta_{21}^*(x_1)$, where $\delta_{21}^*(x_1)$ is the unique $\delta \in (0, 1]$ that satisfies $\delta f(\frac{\delta}{2-\delta}x_1) = \delta^2 \frac{f(x_1)+b_2}{2}$, and δ_{22}^* is defined shortly.

I now show that when $\delta \in (0, \delta_2^*)$, the only SPE of the alternating-offer arbitration game is that at Stage 1, Player 1 offers $(x_1^*(\delta), y_1^*(\delta))$, and Player 2 accepts the offer

[29] Using tie-breaking rule 2.

immediately.

Suppose $\delta \in (0, \delta_2^*)$. Suppose Player 1 makes an offer $(x_1, y_1) \in PF$ at Stage 1. There are two cases:

Case 1: $x_1 = b_1$.

In this case, Player 2 rejects Player 1's offer $(b_1, 0)$ and offers $(\hat{x}_2(b_1,0), f(\hat{x}_2(b_1,0)))$ at Stage 2. This is because $\delta f(\frac{\delta}{2-\delta}b_1) > \delta^2 \frac{b_2}{2} > 0$, where the first inequality follows from the fact that $\delta < \delta_2^* \leq \delta_{13}^*$. Thus, if $\delta \in (0, \delta_2^*)$, then Player 1's equilibrium payoff from offering $(b_1, 0)$ must be $\delta \frac{\delta}{2-\delta} b_1 = \frac{\delta^2}{2-\delta} b_1$.

Case 2: $x_1 < b_1$.

If $0 < \delta < \delta_{21}^*$, then $\delta f(\frac{\delta}{2-\delta}x_1) > \delta^2 \frac{f(x_1)+b_2}{2}$ for any $x_1 \in [0, b_1)$. That is, we must have $B \succ_2 C$. Now by Lemma 16, we must have that $A \sim_2 B \succ_2 C$. $A \sim_2 B$ implies that $f(x_1) = \delta f(\frac{\delta}{2-\delta}x_1)$. So, we must have $x_1 = x_1^*(\delta)$.

Now, let us compare the above two cases. Note that $x_1^*(\delta) \to b_1$ as $\delta \to 0$. Thus, there exists a $\delta_{22}^* \in (0,1)$ such that if $\delta \in (0, \delta_{22}^*)$, then $x_1^*(\delta) > \frac{\delta^2}{2-\delta} b_1$, i.e., for Player 1, the offer $(b_1, 0)$ is dominated by the offer $(x_1^*(\delta), f(x_1^*(\delta)))$.

In summary, if $\delta \in (0, \delta_2^*)$ where $\delta_2^* = \min\{\delta_{13}^*, \delta_{21}^*, \delta_{22}^*\}$, the only SPE of the alternating-offer arbitration game is that Player 1 offers $(x_1^*(\delta), y_1^*(\delta))$, which Player 2 accepts immediately. □

1.6.3 Appendix 3: Other proofs

Sketch of proof of Theorem 6:

To solve for the SPE of the game, note that if Player 1 offers $(x_1, y_1) \in PF$ at Stage 1, then Player 2 either accepts it, or rejects it with one of two counteroffers:

$(\frac{\delta}{2-\delta}x_1, 1-\frac{\delta}{2-\delta}x_1)$ that Player 1 accepts, or $(0,1)$ that Player 1 rejects. The corresponding payoffs are x_1, $\frac{\delta^2}{2-\delta}x_1$ and $\delta^2\frac{x_1}{2}$ for Player 1, and $1-x_1$, $\delta(1-\frac{\delta}{2-\delta}x_1)$ and $\delta^2\frac{2-x_1}{2}$ for Player 2. Player 2 selects the action that maximizes his payoff and the remaining calculations are straightforward. □

Proof of the statement that $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0,1)$ and $\delta x^R(\delta)$ is strictly increasing in $\delta \in (0,1)$:

I first show that $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0,1)$. It is sufficient to show that $x^R(\delta)$ is strictly decreasing in $\delta \in (0,1)$.

Notice that $x^R(\delta)$ satisfies $\delta f(\delta x^R(\delta)) = f(x^R(\delta))$. Differentiating with respect to δ and rearranging terms yields:

$$(f'(x^R(\delta)) - \delta^2 f'(\delta x^R(\delta)))x^{R'} = f(\delta x^R(\delta)) + \delta x^R(\delta) f'(\delta x^R(\delta)). \tag{1}$$

Using the fact that f is concave and strictly decreasing, we have:

$$f'(x^R(\delta)) - \delta^2 f'(\delta x^R(\delta)) < f'(x^R(\delta)) - f'(\delta x^R(\delta)) \leq 0. \tag{2}$$

Notice that the Nash bargaining solution $(x^N, f(x^N))$ satisfies $f'(x^N) = -\frac{f(x^N)}{x^N}$, i.e., $f(x^N) + f'(x^N)x^N = 0$. Since $\delta x^R f(\delta x^R) = x^R f(x^R)$, the two points $(\delta x^R(\delta), f(\delta x^R(\delta)))$ and $(x^R(\delta), f(x^R(\delta)))$ must lie on the curve $xy = c$ with the same constant c. Thus, it must be true that $\delta x^R(\delta) < x^N$ for any $\delta \in (0,1)$. Then,

$$f(\delta x^R(\delta)) + \delta x^R(\delta) f'(\delta x^R(\delta)) > f(x^N) + f'(x^N)x^N = 0 \tag{3}$$

where the inequality follows from the facts that $\delta x^R(\delta) < x^N$ and that f is a concave and strictly decreasing function.

Now, using equations 1, 2, 3, we have $x^{R'} < 0$, i.e., x^R is strictly decreasing in $\delta \in (0, 1)$. We thus proved that $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0, 1)$.

Finally, $\delta x^R(\delta)$ is strictly increasing in $\delta \in (0, 1)$. This follows from the following two facts: (i) the point $(\delta x^R(\delta), f(\delta x^R(\delta)))$ and the point $(x^R(\delta), f(x^R(\delta)))$ lie on the same "indifference curve" $xy = c$, and (ii) $x^R(\delta)$ is strictly decreasing in $\delta \in (0, 1)$. $\square$

1.6.4 Appendix 4: Additional results

In this appendix, I present additional results regarding the Nash final-offer arbitration game and the KS final-offer arbitration game.

Appendix 4.1: Robustness of the Nash final-offer arbitration rule

Yildiz (2011) shows that the unique SPE outcome of the arbitration game that uses the Nash final-offer arbitration rule is $(x^R, f(x^R))$. Here, we provide a simple proof of this result based on Theorem 1.

Theorem 16. *(Yildiz 2011) In the alternating-offer arbitration game where h is the Nash final-offer arbitration rule, the unique SPE outcome is $(x^R, f(x^R))$ for any given $\delta < 1$.*

Proof: Given that $x^R f(x^R) = \delta x^R f(\delta x^R)$, we have $g(x^R) = \delta x^R$. That is, $g(f^{-1}(f(x^R))) = \delta x^R$, which implies that the point $(\delta x^R, f(x^R))$ is on the arbitration curve. Since $(\delta x^R, f(x^R))$ is also on the discounted Pareto frontier, $(\delta x^R, f(x^R))$ is exactly the intersection point of the arbitration curve and the discounted Pareto frontier and we must have $\hat{x}(\delta) = \delta x^R$. Thus, the Nash final-offer arbitration rule is balanced for any discount factor. According to Theorem 1 (i), the unique SPE outcome of the arbitration game that uses the Nash final-offer arbitration rule is $(x^R, f(x^R))$. $\square$

Although for a *given* δ, there exists a class of final-offer arbitration rules under which the arbitration game yields an SPE outcome of $(x^R, f(x^R))$, the following theorem shows that if we require the arbitration game to yield an SPE outcome of $(x^R, f(x^R))$ for *any* $\delta < 1$, then we must have $(x^*, f(x^*)) = (x^N, f(x^N))$ where $(x^N, f(x^N)) = \arg\max\{xy|(x, y) \in S\}$. That is, the arbitrator's ideal settlement must be the Nash solution outcome.

Theorem 17. *In the arbitration game where $h \in \Sigma$, if the SPE outcome is $(x^R, f(x^R))$ for any $\delta < 1$, then we must have $(x^*, f(x^*)) = (x^N, f(x^N))$.*

Sketch of Proof: The theorem follows from the following facts: (i) as δ approaches 1, both δx^R and $\frac{1}{\delta}x^R$ converge to x^N (see Binmore et al. (1986)); (ii) as δ approaches 1, $\hat{x}$ converges to x^*; and (iii) if the SPE outcome of the arbitration game is $(x^R, f(x^R))$ for any $\delta < 1$, then we must have $\delta x^R \leq \hat{x} \leq \frac{1}{\delta}x^R$ for any $\delta < 1$. $\square$

Appendix 4.2: Additional results for the KS final-offer arbitration game

Theorem 18. *Suppose the final-offer arbitration rule is such that $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ if and only if $u^{KS}(x_2, y_2) \geq u^{KS}(x_1, y_1)$ (u^{KS} is defined in Section 3.3). Then,*

(i) if $\dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R) f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R} \leq \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$, then the unique SPE outcome of the final-offer arbitration game is $(x^R, f(x^R))$.

(ii) if $\dfrac{b_2}{b_1} < \dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R) f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R}$*, then the unique SPE of the final-offer arbitration game is a type-II arbitration-driven equilibrium. That is, at Stage 1, Player 1 makes the offer* $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$ *that Player 2 rejects, and at Stage 2, Player 2 makes the counteroffer* $(\hat{x}, f(\hat{x}))$ *that Player 1 accepts. Moreover,* $\hat{x}$ *is determined by the equality that* $\dfrac{\hat{x}/f(\hat{x})}{b_1/b_2} = \dfrac{f(\frac{1}{\delta}\hat{x})/\frac{1}{\delta}\hat{x}}{b_2/b_1}$.

(iii) if $\dfrac{b_2}{b_1} > \dfrac{f(x^R)}{\delta x^R}$*, then the unique SPE of the final-offer arbitration game is a type-I arbitration-driven equilibrium. That is, Player 1 offers* $(f^{-1}(\delta f(\hat{x})), \delta f(\hat{x}))$ *that Player 2 accepts. Moreover,* $\hat{x}$ *is determined by the equality that* $\dfrac{\hat{x}/f(\hat{x})}{b_1/b_2} = \dfrac{\delta f(\hat{x})/f^{-1}(\delta f(\hat{x}))}{b_2/b_1}$.

Sketch of proof: First, $\hat{x} \geq \delta x^R$ is equivalent to $g(x^R) \geq \delta x^R$. The condition $g(x^R) \geq \delta x^R$ (i.e., the arbitrator weakly prefers $(x^R, f(x^R))$ over $(\delta x^R, f(\delta x^R))$) is satisfied if and only if (i) $\dfrac{f(x^R)/x^R}{b_2/b_1} < 1$, $\dfrac{f(\delta x^R)/\delta x^R}{b_2/b_1} > 1$, and $\dfrac{f(x^R)/x^R}{b_2/b_1} \geq (\dfrac{f(\delta x^R)/\delta x^R}{b_2/b_1})^{-1}$, or (ii) $\dfrac{f(x^R)/x^R}{b_2/b_1} \geq 1$. The condition (i) is equivalent to $\dfrac{b_2}{b_1} > \dfrac{f(x^R)}{x^R}$, $\dfrac{b_2}{b_1} < \dfrac{f(\delta x^R)}{\delta x^R}$, and $\dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$ (for the last inequality, using the fact that $\delta f(\delta x^R) = f(x^R)$), which is equivalent to $\dfrac{f(x^R)}{x^R} < \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$. The condition (ii) is equivalent to $\dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{x^R}$. Thus, the condition (i) or (ii) is satisfied if and only if $\dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$.

Second, $\hat{x} \leq \frac{1}{\delta}x^R$ is equivalent to $g(\dfrac{1}{\delta^2}x^R) \leq \dfrac{1}{\delta}x^R$. The condition $g(\dfrac{1}{\delta^2}x^R) \leq \dfrac{1}{\delta}x^R$ is equivalent to $\dfrac{b_2}{b_1} \geq \dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R) f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R}$ (the proof here is similar to the first step and is thus omitted). $\square$

Theorem 18 then follows from the above analysis and Theorem 1.

Now, we can prove Theorem 4.

Proof of Theorem 4:

By Theorem 18, if $\frac{\sqrt{\delta f(\frac{1}{\delta^2}x^R)f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R} \leq \frac{b_2}{b_1} \leq \frac{f(x^R)}{\delta x^R}$, then the unique SPE outcome of the KS final-offer arbitration game is $(x^R, f(x^R))$. A sufficient condition for $\frac{\sqrt{\delta f(\frac{1}{\delta^2}x^R)f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R} \leq \frac{b_2}{b_1} \leq \frac{f(x^R)}{\delta x^R}$ is $\frac{f(\frac{1}{\delta}x^R)}{\frac{1}{\delta}x^R} \leq \frac{b_2}{b_1} \leq \frac{f(x^R)}{\delta x^R}$. This follows from the fact that $f(\frac{1}{\delta^2}x^R) \leq f(\frac{1}{\delta}x^R)$ and $\delta < 1$. $\square$

1.7 Chapter 1 References

N. Anbarci, J. H. Boyd, *Nash demand game and the Kalai-Smorodinsky solution*, Games and Economic Behavior 71 (2011), 14-22.

K. Binmore, A. Rubinstein, A. Wolinsky, *The Nash bargaining solution in economic modelling*, Rand Journal of Economics 17 (1986), 176-188.

K. Binmore, A. Shaked, J. Sutton, *An outside option experiment*, Quarterly Journal of Economics 104 (1989), 753-770.

J. R. Chelius, J. B. Dworkin, *An economic analysis of final-offer arbitration as a conflict resolution device*, Journal of Conflict Resolution 24 (1980), 293-310.

O. Compte, P. Jehiel, *On the role of arbitration in negotiations*, mimeo C.E.R.A.S.-E.N.P.C. (1995).

C. A. Deck, A. Farmer, *Bargaining over an uncertain value: arbitration mechanisms compared*, Journal of Law, Economics, and Organization 23 (2007), 547-579.

P. Feuille, *Final offer arbitration and the chilling effect*, Industrial Relations 14 (1975), 302-310.

E. Kalai, M. Smorodinsky, *Other solutions to Nash's bargaining problem*, Econometrica 43 (1975), 513-518.

P. Manzini, M. Mariotti, *Perfect equilibria in a model of bargaining with arbitration*, Games and Economic Behavior 37 (2001), 170-195.

P. Manzini, M. Mariotti, *Going alone together: joint outside options in bilateral negotiations*, Economic Journal 114 (2004), 943-960.

J. Nash, *The bargaining problem*, Econometrica 18 (1950), 155-162.

C. Ponsatí, J. Sákovics, *Rubinstein bargaining with two-sided outside options*, Economic Theory 11 (1998), 667-672.

K. Rong, *An axiomatic approach to arbitration and its application in bargaining games*, Working paper (2011).

A. Rubinstein, *Perfect equilibrium in a bargaining model*, Econometrica 50 (1982), 97-110.

A. Shaked, *Opting out: bazaars versus "hi tech" markets*, Investigaciones Económics 18 (1994), 421-432.

C. Stevens, *Is compulsory arbitration compatible with bargaining?* Industrial Relations 5 (1966), 38-52.

R. Wilson, *Negotiation with private information: litigation and strikes*, Nancy L Schwartz Memorial Lecture (1994), Northwestern University.

M. Yildiz, *Nash meets Rubinstein in final-offer arbitration*, Economics Letters 110 (2011), 226-230.

Chapter 2

An Axiomatic Approach to Arbitration and Its Application in Bargaining Games

2.1 Introduction

Arbitration occurs when two players are unable to reach agreement with each other. In this paper, we formally define the arbitration problem as the triplet that consists of offers submitted by two players and their bargaining set. An arbitration solution outcome is a point in the bargaining set chosen by an arbitrator. In order to obtain the arbitration outcome, the arbitrator usually follows a certain arbitration procedure. In the literature, there are two well-know arbitration procedures. One is the rule of equally-split-the-difference between players' offers, and the other is the final-offer arbitration rule.[30]

In this paper, we will use the axiomatic approach (Nash, 1950; Kalai and Smorodinsky, 1975) to characterize the arbitration procedure. An advantage of the axiomatic approach is that, we don't need to characterize the detailed arbitration process. Instead, we propose several axioms that an arbitration procedure should

[30]Final offer arbitration is a procedure in which the arbitrator must choose one of the players' offers as the arbitration outcome (Stevens, 1966).

satisfy and then find the arbitration solution that satisfies those axioms.

The key axiom we impose on the arbitration procedure is "Symmetry in Offers," which requires *fairness* in arbitration. More particularly, it requires that whenever the two players' offers are symmetric with each other, the arbitrated outcome should also be symmetric. "Symmetry in Offers" appears to be a strong rule in the sense that it does not require symmetry in the bargaining set. However, "Symmetry in Offers" is a natural rule given that the arbitrator should make a decision primarily based on players' offers, instead of the shape of the bargaining set. In addition, it is a simple rule because it does not require the arbitrator to calculate the entire shape of the bargaining set. The other two axioms, "Invariance w.r.t Affine Transformation" and "Pareto Optimality" are self-evident. They require *invariance* and *efficiency* in arbitration respectively. We find that there is a unique arbitration solution that satisfies all the three axioms. We call this solution the *symmetric arbitration solution*. The symmetric arbitration solution has a simple graphical representation: for any given bargaining set and offers submitted by the two players, the symmetric arbitration solution outcome is the intersection point of the Pareto frontier of the bargaining set with the line joining the component-wise minimum and component-wise maximum of the offers. When the Pareto frontier of the bargaining set is linear, the symmetric arbitration solution coincides with the rule of "equally splitting the difference." The symmetric arbitration solution solution is "superior" to the rule of "equally splitting the difference" in that when the Pareto frontier of the bargaining set is nonlinear, "equally splitting the difference" results in an *inefficient* outcome, while the symmetric arbitration solution results in an *efficient* outcome.

Although our focus is to use "Symmetry in Offers" to characterize the symmetric arbitration solution, it is possible for us to use the weaker version of "Symmetry in Offers" to characterize the symmetric arbitration solution. The weaker version of "Symmetry in Offers," called "Weak Symmetry in Offers," requires that the arbitration solution outcome be symmetric whenever players' offers are symmetric *and* the bargaining set is symmetric. We show that the symmetric arbitration solution is the only solution that satisfies "Weak Symmetry in Offers," "Invariance," "Pareto Optimality," and "Strong Monotonicity."

We then propose two bargaining games in which, whenever the players are unable to reach agreement, an arbitration stage is reached and the *symmetric arbitration solution* is used to decide the outcome. The first game is a simultaneous-offer game. In this game, two players make offers simultaneously. If the offers are compatible, then each player gets what he demands, otherwise the game moves to the arbitration stage. In the arbitration stage, the *symmetric arbitration solution* is utilized to determine the outcome. This game is similar to the second Nash demand game in Anbarci and Boyd (2011)[31]. Both games are variants of the Nash demand game (Nash, 1953) and have arbitration stages. The difference is that the game in Anbarci and Boyd (2011) uses the rule of "equally splitting the difference" at the arbitration stage, but our game uses the symmetric arbitration solution.

[31] The game considered in Anbarci and Boyd (2011) can be rephrased as follows. At the first stage, two players make offers simultaneously. If the offers are compatible, then each player gets what he demands, otherwise with probability $1 - p$ the game terminates with the disagreement point as the outcome, and with probability p the game goes to the arbitration stage in which the rule of "equally splitting the difference" is used to decide the outcome. Note that the probability p of moving to the arbitration stage in their game is equivalent to the discount factor δ in our game when the disagreement point is normalized to $(0, 0)$.

Our second game is an alternating-offer game. In this game, at stage 1, player 1 makes an offer and player 2 decides whether to accept or reject it. If player 2 chooses to reject the offer, then the game moves to the next stage, at which player 2 makes an offer and player 1 decides whether to accept or reject it. If player 1 rejects the offer, then the game moves to the arbitration stage in which the *symmetric arbitration solution* is used to decide the final outcome. This game can be regarded as a variant of the game proposed by Yildiz (2011) and the two games studied by Rong (2011). In all of those games, two players make offers sequentially and if both offers are rejected, the game moves to an arbitration stage. Our game differs from Yildiz (2011) and Rong (2011) in that our game uses the symmetric arbitration solution at the arbitration stage, while the game in Yildiz (2011) and the first game in Rong (2011) use final offer arbitration, and the second game in Rong (2011) uses equally-split-the-difference arbitration.

In both the simultaneous-offer game and the alternating-offer game that we consider, the only arbitration cost is the time cost, which is measured by the common discount factor of players. Our equilibrium analyses show that, in both games, when the discount factor is close to 1 (i.e., the time cost is low), players tend to make extreme offers. The threshold discount factor required for players to make extreme offers is relatively small. In particular, when the Pareto frontier is linear, the threshold discount factor is $\frac{2}{3}$ for the simultaneous-offer game and is 0.91 for the alternating-offer game. In addition, we find that, when both players make extreme offers, the arbitrated outcome coincides with the Kalai-Smorodinsky solution outcome.

The result that as the discount factor becomes close to 1, the only equilibrium requires each player to make the extreme offer is not surprising. Actually, it is well known in the literature that if a bargaining process involves an arbitration mechanism which allows for compromise between offers, then the bargaining process is subject to the so-called *chilling effect* (Feuille, 1975; Deck and Farmer, 2007). That is, players tend to take extreme positions before arbitration. This tendency is stronger when players become more patient.

This paper is organized as follows. The next subsection is the axiomatic characterization of the arbitration problem. Section 2.3 presents the main result. Section 2.4 provides an alternative axiomatic characterization of the symmetric arbitration solution using the axiom of Weak Symmetry in Offers. Section 2.5 discusses the two bargaining games with symmetric arbitration, i.e., the "simultaneous-offer game with symmetric arbitration" and the "alternating-offer game with symmetric arbitration." Concluding remarks are offered in section 2.6.

2.2 Axiomatic Characterization of Arbitration Problem

Suppose there are two players who are expected utility maximizers. Let $S \subset R^2$ denote the bargaining set, which includes all possible bargaining outcomes, measured in expected utility level. Let $(x_1, y_1) \in S$ denote player 1's final offer submitted to an arbitrator and $(x_2, y_2) \in S$ denote player 2's final offer submitted to the arbitrator. Note we always use x to represent player 1's payoff and y to represent player 2's payoff.

We assume the bargaining set S is nonempty, convex, compact and strictly comprehensive. The definition of "comprehensiveness" and "strict comprehensiveness" are given below:

Definition 2. *S is **comprehensive** if $\exists(d_1, d_2) \in R^2$ s.t. $\forall(x, y) \in S$, we have (i) $(x, y) \geq (d_1, d_2)$, and (ii) if $(d_1, d_2) \leq (x', y') \leq (x, y)$, then $(x', y') \in S$.*

Definition 3. *S is **strictly comprehensive** if S is comprehensive and for any $(x, y) \in S$ and $(x', y') \in S$ with $(x', y') \geq (x, y)$ and $(x', y') \neq (x, y)$, there exists a $(x'', y'') \in S$ such that $(x'', y'') >> (x, y)$.*

If we regard d as the disagreement point, then the "comprehensiveness" of a bargaining set simply requires: (i) for each player, the utility level at the disagreement point is the lowest possible utility level that he can get from bargaining; (ii) each player can freely dispose any utility that is higher than the disagreement point.

Strict comprehensiveness further requires the Pareto frontier of the bargaining set be strictly downward-sloping. We need a bargaining set to be strictly comprehensive to avoid the case that the Pareto frontier contains a flat or vertical segment. A typical strictly comprehensive bargaining set S is shown in Figure 15.

Any nonempty bargaining set S that is convex, compact and strictly comprehensive determines a *unique* $d = (d_1, d_2)$ that satisfies Definition 1. We use $d(S)$ to denote this point.

The Pareto frontier of the bargaining set S is defined as $PF(S) = \{p \in S : q \geq p$ with $q \neq p \Rightarrow q \notin S\}$. We assume that each player can only make an offer *on* the Pareto frontier. This assumption is made for simplicity, although it is not essential for our main results.

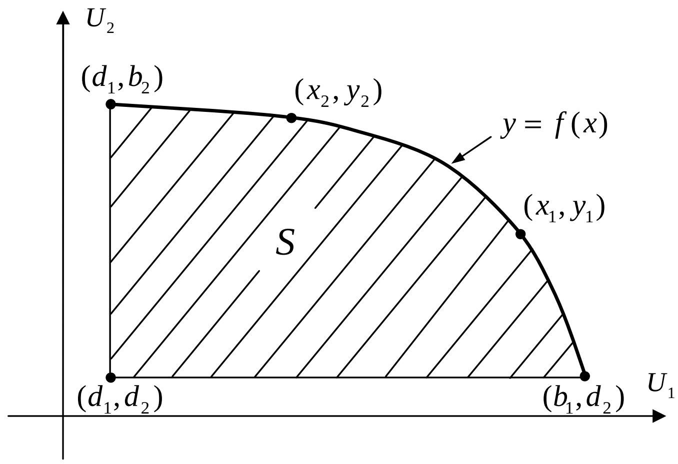

Figure 15: Bargaining set and players' offers.

Define $b_i = max\{U_i : (U_1, U_2) \in S\}$ as player i's maximal possible utility level from the bargaining set. Define the function $f : x \to max\{y|(x, y) \in S\}$ for $x \in [d_1, b_1]$. Thus $\{(x, f(x))|x \in [d_1, b_1]\}$ denotes the Pareto frontier. Our assumption that S is convex, compact and strictly comprehensive implies that f is a strictly decreasing function on $[d_1, b_1]$ with $f(d_1) = b_2$ and $f(b_1) = d_2$.

Now define $\Sigma = \{S \subset R^2 \,|S$ is nonempty, convex, compact, strictly comprehensive$\}$ and $\mathcal{B} = \{((x_1, y_1), (x_2, y_2), S) \,|(x_1, y_1) \in PF(S)\,,\ (x_2, y_2) \in PF(S),\ (x_1, y_2) \notin S$ and $S \in \Sigma\}$. We call any $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$ an *arbitration problem*.[32] An *arbitration solution* is any function $g : \mathcal{B} \to R^2$ such that $g((x_1, y_1), (x_2, y_2), S) \in S$. We may write $g((x_1, y_1), (x_2, y_2), S) = (g_1((x_1, y_1), (x_2, y_2), S), g_2((x_1, y_1), (x_2, y_2), S))$, where $g_i((x_1, y_1), (x_2, y_2), S)$ is the arbitration outcome for player i.

[32]Notice that the arbitration problem we consider involves *incompatible* offers (i.e., $(x_1, y_2) \notin S$). If players' offers are compatible, then each player simply gets what he demands (and arbitration is not necessary). In addition, notice that our arbitration problem consists of two players' offers and a bargaining set, while the classic bargaining problem proposed by Nash (1950) consists of a disagreement point and a bargaining set.

We will propose the following three axioms that an arbitration solution should satisfy:

Definition 4. *An arbitration solution g is a* ***symmetric arbitration solution*** *if it satisfies the following three axioms:*

1. *Axiom 1 (Symmetry in Offers): For any arbitration problem $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$ with $x_1 = y_2$ and $x_2 = y_1$, we have $g_1((x_1, y_1), (x_2, y_2), S) = g_2((x_1, y_1), (x_2, y_2), S)$.*

2. *Axiom 2 (Invariance w.r.t. Affine Transformation): If $A : R^2 \to R^2$ represents a strictly increasing affine transformation, i.e., $A(x, y) = (a_1 x + c_1, a_2 y + c_2)$ for some positive constant a_i and some constant c_i, then we have $g(A(x_1, y_1), A(x_2, y_2), A(S)) = A(g((x_1, y_1), (x_2, y_2), S))$ for any $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$.*

3. *Axiom 3 (Pareto Optimality): For any arbitration problem $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$, we have $g((x_1, y_1), (x_2, y_2), S) \in PF(S)$.*

Axiom 1 requires that if the offers from the two players are symmetric around the 45 degree line, then the arbitration solution outcome should also be symmetric (i.e., on the 45 degree line). That is, if each player makes the same demand for himself and suggests the same payoff for his opponent, then the arbitrated outcome should result in the same payoff for each player. Axiom 1 does not require symmetry in the bargaining set. However, we still regard Axiom 1 as a natural rule for the following reasons. First, an arbitrator should primarily focus on the offers of players, instead of the shape of the bargaining set. Second, it is generally costly for the arbitrator to calculate the entire shape of the bargaining set. Axiom 1 (together with Axiom 2 and Axiom 3) only requires that the arbitrator calculate a fraction of the bargaining set in order to determine the arbitration outcome on the Pareto frontier.[33]

Axiom 2 is adapted from Nash (1950). The idea behind this axiom is that the arbitration outcome should only depend on players' underlying preferences and not on their utility representations. Hence, for two arbitration problems with the same preferences and the same physical offers submitted by the players, the arbitration outcome should also be the same (with correspondingly different utility representation). Note that players' utilities are expected utilities, so a player's utility is unique up to strictly increasing affine transformation. Finally, Axiom 3 simply requires the arbitration outcome to be efficient.

2.3 Main Result

It turns out the symmetric arbitration solution is unique and has a simple representation. For $p_1, p_2 \in R^2$, let $L(p_1, p_2)$ denote the line joining p_1 and p_2. We have the following result:

Theorem 5. *There is one and only one symmetric arbitration solution, denoted by γ. The function γ has the following simple graphic representation. For any arbitration problem $((x_1, y_1), (x_1, y_2), S) \in \mathcal{B}$, $\gamma((x_1, y_1), (x_2, y_2), S)$ is the intersection point of $L((x_1, y_1) \wedge (x_2, y_2), (x_1, y_1) \vee (x_2, y_2))$ with $PF(S)$ (see Figure 16).*

Proof: For a given arbitration problem $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$, we have two

[33] I am indebted to an editor and an anonymous referee for suggesting the above explanations for the axiom of Symmetry in Offers.

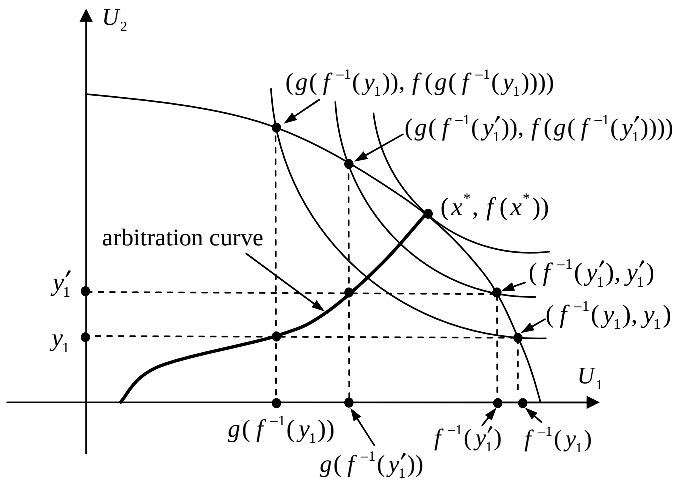

Figure 16: Symmetric arbitration solution.

cases:

(i) $y_2 > x_2$.

We need to find a strictly increasing affine transformation that transforms the given problem $((x_1, y_1), (x_2, y_2), S)$ to an offer-symmetric problem $((x'_1, y'_1), (x_2, y_2), S')$, where $x'_1 = y_2$ and $y'_1 = x_2$ (see Figure 17). Let $A^*_i(x) = a^*_i x + c^*_i$ $(i = 1, 2)$ be such a transformation. Then we have:

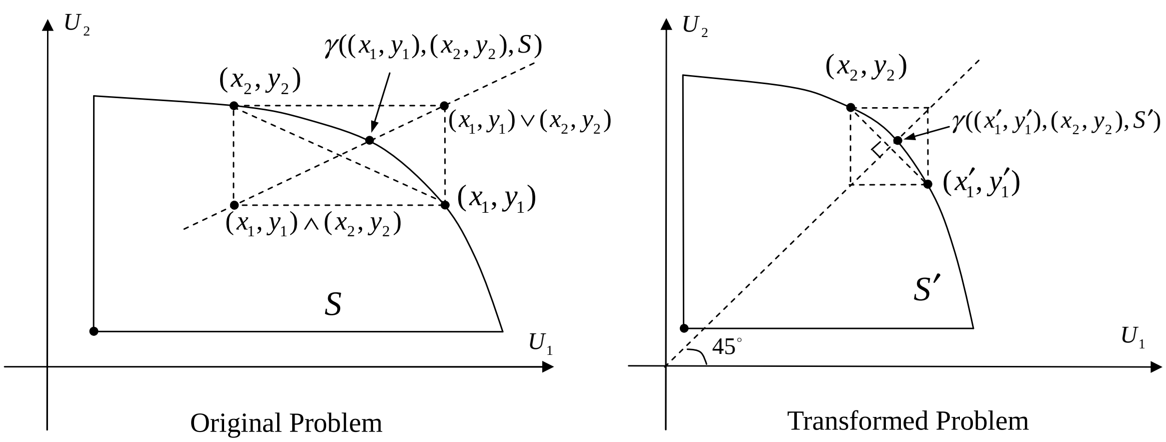

Figure 17: Transformation of the arbitration problem.

$$\begin{cases} x_2 = a_1^* x_2 + c_1^* \\ y_2 = a_2^* y_2 + c_2^* \end{cases} \text{and} \begin{cases} y_2 = a_1^* x_1 + c_1^* \\ x_2 = a_2^* y_1 + c_2^* \end{cases}. \tag{4}$$

Solving the equations, we have:

$$\begin{cases} a_1^* = \dfrac{y_2 - x_2}{x_1 - x_2} \\ a_2^* = \dfrac{x_2 - y_2}{y_1 - y_2} \end{cases} \text{and} \begin{cases} c_1^* = \dfrac{x_2(x_1 - y_2)}{x_1 - x_2} \\ c_2^* = \dfrac{y_2(y_1 - x_2)}{y_1 - y_2} \end{cases}. \tag{5}$$

Since $(x_1, y_1) \in PF(S)$, $(x_2, y_2) \in PF(S)$, $(x_1, y_2) \notin S$ and the Pareto frontier is strictly downward-sloping, we must have $x_1 > x_2$ and $y_1 < y_2$. Note we have also assumed that $y_2 > x_2$. It can be verified that $a_1^* > 0$ and $a_2^* > 0$, which ensures that the above affine transformation is indeed an expected utility transformation.

If (u_1^*, u_2^*) is the symmetric arbitration solution to the original arbitration problem $((x_1, y_1), (x_2, y_2), S)$, then by Axiom 2, $(a_1^* u_1^* + c_1^*, a_2^* u_2^* + c_2^*)$ is the symmetric arbitration solution to the transformed problem $((x_1', y_1'), (x_2, y_2), S')$. Since $((x_1', y_1'), (x_2, y_2), S')$ is symmetric in offers, the symmetric arbitration solution to it must be on the 45 degree line. Hence, we have:

$$a_1^* u_1^* + c_1^* = a_2^* u_2^* + c_2^*. \tag{6}$$

Using equations 5, equation 6 can be rewritten as:

$$u_2^* = \frac{y_2 - y_1}{x_1 - x_2} u_1^* + \frac{x_1 y_1 - x_2 y_2}{x_1 - x_2}. \tag{7}$$

It can be verified that the line

$$u_2 = \frac{y_2 - y_1}{x_1 - x_2} u_1 + \frac{x_1 y_1 - x_2 y_2}{x_1 - x_2}$$

is the line that passes through $(x_1, y_1) \wedge (x_2, y_2)$ and $(x_1, y_1) \vee (x_2, y_2)$. Now, by Axiom 3 (Pareto Optimality), we can conclude that (u_1^*, u_2^*) must be the intersection point of $L((x_1, y_1) \wedge (x_2, y_2), (x_1, y_1) \vee (x_2, y_2))$ with the Pareto frontier.

(ii) $y_2 \leq x_2$***.*** We can always find a strictly increasing affine transformation such that the transformed arbitration problem has the property $y_2' > x_2'$. Then we go back to case (i) and the remaining proof is straightforward. $\square$

A graphic representation of the symmetric arbitration solution is shown in Figure 16.

The idea of the proof is that, for any offer-nonsymmetric arbitration problem $((x_1, y_1), (x_2, y_2), S)$, we can always find a strictly increasing affine transformation to transform it to an offer-symmetric problem $((x_1', y_1'), (x_2, y_2), S')$, where $x_1' = y_2$ and $y_1' = x_2$ (see Figure 17). Due to the axiom of Pareto optimality and the axiom of Symmetry in Offers, the symmetric arbitration solution to the problem $((x_1', y_1'), (x_2, y_2), S')$ must be the intersection point of the 45 degree line with the Pareto frontier. Then, using the inverse of the above affine transformation, we can transform this solution outcome back to the original problem. It can be verified that the solution to the original problem is exactly the intersection point of $L((x_1, y_1) \wedge (x_2, y_2), (x_1, y_1) \vee (x_2, y_2))$ with $PF(S)$.

Another graphic interpretation of the solution is as follows. For the arbitration

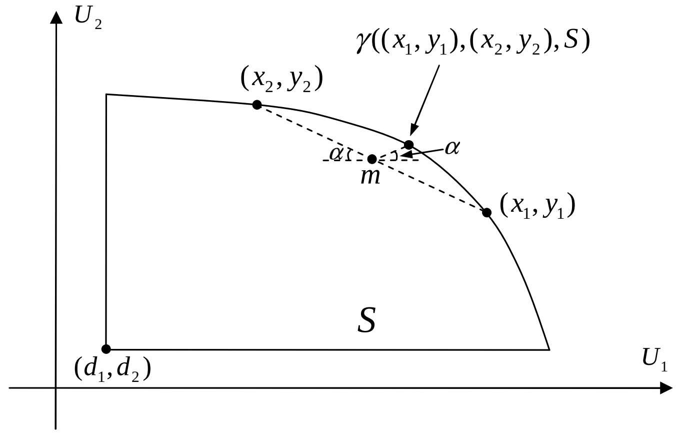

Figure 18: Another representation of the symmetric arbitration solution.

problem $((x_1, y_1), (x_2, y_2), S)$, connect the two offers (x_1, y_1) and (x_2, y_2) with a line and denote its middle point by m. Now, draw a line through m with a slope that is the negative of the slope of $L((x_1, y_1), (x_2, y_2))$. Then, the intersection point of this new line with the Pareto frontier is the symmetric arbitration solution (see Figure 18). The essential point here is that the line joining m and the solution point $\gamma((x_1, y_1), (x_2, y_2), S)$ always has a slope that is the negative of the slope of the line joining (x_1, y_1) and (x_2, y_2). Note this is true for any offer-symmetric arbitration problem because of the axiom of Symmetry in Offers. This is also true for any offer-nonsymmetric arbitration problem, because (i) any offer-nonsymmetric problem can be transformed to an offer-symmetric problem by some strictly increasing affine transformation, and (ii) two lines with slopes that are opposite in sign is a property preserved by any affine transformation.[34]

When the Pareto frontier of the bargaining set is linear, the symmetric arbitra-

[34]See Nash (1953) for a similar geometric explanation for the Nash bargaining solution.

tion solution outcome coincides with the outcome of "equally splitting the difference." When the Pareto frontier of the bargaining set is nonlinear, "equally splitting the difference" results in an *inefficient* outcome (point m in Figure 18), while the symmetric arbitration solution results in an *efficient* outcome.

2.4 Another Axiomatic Characterization of Symmetric Arbitration Solution[35]

In this section, we propose a weaker version of the axiom of Symmetry in Offers, called Weak Symmetry in Offers. It requires that the arbitration solution outcome be symmetric whenever players' offers are symmetric *and* the bargaining set is symmetric. It turns out that the symmetric arbitration solution is the unique arbitration solution that satisfies the following four axioms: Weak Symmetry in Offers, Invariance, Pareto Optimality, and Strong Monotonicity.

Definition 6. *Let g be an arbitration solution. The axiom of Weak Symmetry in Offers and the axiom of Strong Monotonicity are defined as follows:*

1. *Axiom $1'$ (Weak Symmetry in Offers): For any arbitration problem $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$ where $x_1 = y_2$, $x_2 = y_1$ and S is symmetric, we have $g_1((x_1, y_1), (x_2, y_2), S) = g_2((x_1, y_1), (x_2, y_2), S)$.*

2. *Axiom 4 (Strong Monotonicity): For any two arbitration problems $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$ and $((x_1, y_1), (x_2, y_2), S') \in \mathcal{B}$, if*

[35] I am indebted to an anonymous referee who suggested that I use a weaker symmetry axiom and some type of monotonicity axiom to characterize the symmetric arbitration solution.

$S' \supset S$, then $g_1(((x_1, y_1), (x_2, y_2), S')) \geq g_1(((x_1, y_1), (x_2, y_2), S))$ and $g_2(((x_1, y_1), (x_2, y_2), S')) \geq g_2(((x_1, y_1), (x_2, y_2), S))$.

Theorem 7. *The symmetric arbitration solution γ is the unique arbitration solution that satisfies Axiom 1′, Axiom 2, Axiom 3, and Axiom 4.*

Proof: It is easy to verify that the symmetric arbitration solution γ satisfies Axiom 1′, Axiom 2, Axiom 3, and Axiom 4. Now, assume that there is another arbitration solution μ that satisfies all the four axioms. We will show that $\mu((x_1, y_1), (x_2, y_2), S) = \gamma((x_1, y_1), (x_2, y_2), S)$ for any $((x_1, y_1), (x_2, y_2), S) \in \mathcal{B}$.

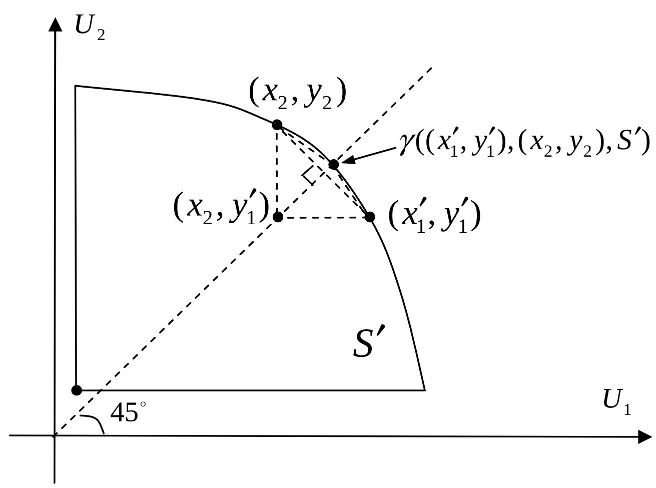

Figure 19: Transformed arbitration problem.

It is without loss of generality to assume that $y_2 > x_2$.[36] Similar to part (i) of proof of Theorem 5, we can find a strictly increasing affine transformation (A_1^*, A_2^*) such that $((x_1, y_1), (x_2, y_2), S)$ can be transformed to an offer-symmetric arbitration problem $((x_1', y_1'), (x_2, y_2), S')$, where $x_1' = y_2$ and $y_1' =$

[36]If $y_2 \leq x_2$, then we can always transform the arbitration problem to a new problem, which has the property $y_2' > x_2'$.

x_2. Figure 19 illustrates the transformed arbitration problem. Let $S'' =$ convex hull $\{(x_1', y_1'), (x_2, y_2), (x_2, y_1'), \gamma((x_1', y_1'), (x_2, y_2), S')\}$. Since S'' is symmetric, by Axiom 1$'$, $\mu((x_1', y_1'), (x_2, y_2), S'') = \gamma((x_1', y_1'), (x_2, y_2), S')$. Since $S' \supset S''$, by Axiom 4, we must have $\mu((x_1', y_1'), (x_2, y_2), S') = \mu((x_1', y_1'), (x_2, y_2), S'')$, so, $\mu((x_1', y_1'), (x_2, y_2), S') = \gamma((x_1', y_1'), (x_2, y_2), S')$. Now, we can use the inverse of the transformation (A_1^*, A_2^*) to transform the solution $\mu((x_1', y_1'), (x_2, y_2), S')$ back to the original problem, and we must have $\mu((x_1, y_1), (x_2, y_2), S) = \gamma((x_1, y_1), (x_2, y_2), S)$. $\square$

2.5 Bargaining Games with Symmetric Arbitration

This section will analyze two bargaining games that involve symmetric arbitration. One is the simultaneous-offer game, and the other is the alternating-offer game.

From this point on, we fix the bargaining set S and we will simply write $\gamma((x_1, y_1), (x_2, y_2), S)$ as $\gamma((x_1, y_1), (x_2, y_2))$ whenever there is no confusion. We use $\delta \in (0, 1]$ to denote the discount factor, which means 1 unit of utility at the next stage is equivalent to δ unit of utility at the current stage. Finally, we assume throughout this section that $d(S) = (0, 0)$.

The following lemma states that a player's payoff obtained from the symmetric arbitration solution is strictly increasing in both his own demand and his opponent's suggested payoff for him. This implies that, if a player takes a stronger position (i.e., demand more) before arbitration, then he will get more payoff from arbitration.[37]

Lemma 8. *For $x_1,\ x_2 \in [0, b_1]$, $\gamma_1((x_1, f(x_1)), (x_2, f(x_2)))$ is strictly increasing in*

[37] This is true for any arbitration procedure that allows for compromises between offers.

x_1 and x_2; $\gamma_2((x_1, f(x_1)), (x_2, f(x_2)))$ is strictly decreasing in x_1 and x_2.

Proof: see the appendix. □

2.5.1 Simultaneous-Offer Game

Simultaneous-offer game is also known as the Nash demand game. In the original Nash demand game (Nash 1953), two players make demands (offers) simultaneously. If their demands are compatible, then each player gets what he demands; otherwise, each player gets the disagreement payoff. One disadvantage of the Nash demand game is that it is a one-stage game that does not allow for renegotiation or arbitration. In the literature, many variants of the Nash demand game have been proposed to deal with this problem (e.g., Howard, 1992; Anbarci and Boyd, 2011).[38] Here, we are going to propose a new Nash demand game, in which players move to an arbitration stage whenever their offers are incompatible. In addition, we assume that the symmetric arbitration solution is used at the arbitration stage. In particular, we define the *simultaneous-offer game (Nash demand game) with symmetric arbitration* as follows:

1. Stage 1: player 1 and player 2 submit their offers simultaneously. Let $(x_1, y_1) \in PF(S)$ be the offer submitted by player 1 and $(x_2, y_2) \in PF(S)$ be the offer submitted by player 2. If (x_1, y_1) and (x_2, y_2) are compatible, then (x_1, y_2) is the outcome. Otherwise, the game moves to Stage 2.

[38] Howard (1992) extended the original Nash demand game to a multi-stage game which allows for “renegotiation”. In Anbarci and Boyd (2011), their second Nash demand game introduced an arbitration stage, in which the rule of “equally splitting the difference” is utilized to decide the arbitration outcome.

2. Stage 2: an arbitrator decides the outcome using the *symmetric arbitration solution*, i.e., $\gamma((x_1,y_1),(x_2,y_2))$ is the arbitrated outcome.

Notice that players' payoffs obtained at stage 2 are discounted by δ. So, if the game moves to arbitration, the arbitrated payoffs received by players are $\delta\gamma((x_1,y_1),(x_2,y_2))$. Before characterizing the equilibria in this game, we will make the following definition (refer to Figure 20).

Definition 9. *For any* $(x,y) \in PF(S)$, *define* $\tilde{x}(x) = \gamma_1((b_1,0),(x,f(x)))$ *and* $\tilde{y}(y) = \gamma_2((f^{-1}(y),y),(0,b_2))$.

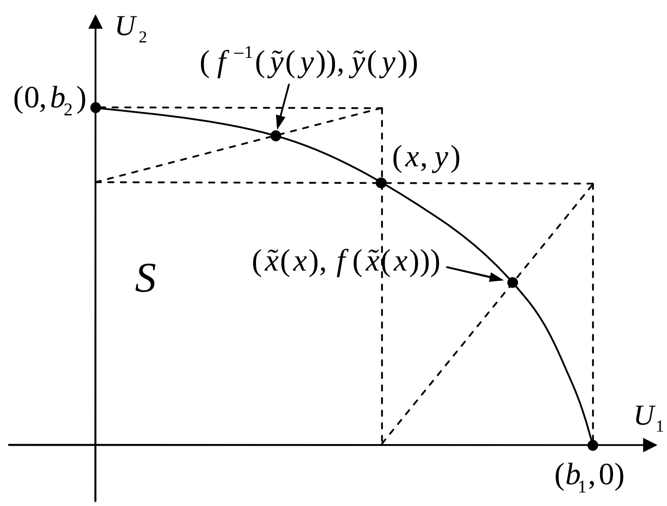

Figure 20: Definition of $\tilde{x}(x)$ and $\tilde{y}(y)$.

$\tilde{x}(x)$ is player 1's stage-2 payoff when his opponent's makes the offer $(x,y) \in PF(S)$ while he makes the extreme offer $(b_1,0)$. On the other hand, $\tilde{y}(y)$ is player 2's stage-2 payoff when his opponent makes the offer $(x,y) \in PF(S)$ while he makes the extreme offer $(0,b_2)$. According to Lemma 8, a player's arbitrated payoff is strictly increasing in his own demand. Thus, $\tilde{x}(x)$ is player 1's best possible (stage-2) arbitrated payoff when his opponent makes the offer (x, y). Similarly, $\tilde{y}(y)$ is player 2's best possible (stage-2) arbitrated payoff when his opponent makes the offer (x, y).

We will use $((x_1, y_1), (x_2, y_2))$ to denote the strategy profile in which player 1 submits the offer (x_1, y_1) and player 2 submits the offer (x_2, y_2). If a player makes the offer (x, y), then the other player can choose to make the same offer (x, y) and obtain x (if he is player 1) or y (if he is player 2), or choose to make the extreme offer (which will move the game to arbitration) and obtain $\tilde{x}(x)$ (if he is player 1) or $\tilde{y}(y)$ (if he is player 2) at the arbitration stage. Thus, $((x, y), (x, y))$ is a Nash equilibrium if and only if $x \geq \delta\tilde{x}(x)$ and $y \geq \delta\tilde{y}(y)$. In addition, $((b_1, 0), (0, b_2))$ is always a Nash equilibrium regardless of how high the discount factor might be.

The following theorem summarizes the results above. It actually describes all the possible Nash equilibria in the simultaneous-offer game with symmetric arbitration.

Theorem 10. *In the simultaneous-offer game with symmetric arbitration, there are two possible types of Nash equilibria:*

(i) (immediate-agreement equilibrium) $((x, y), (x, y))$ $((x, y) \in PF(S))$ *is a Nash equilibrium if and only if* $x \geq \delta\tilde{x}(x)$ *and* $y \geq \delta\tilde{y}(y)$*;*

(ii) (arbitration equilibrium) $((b_1, 0), (0, b_2))$ *is a Nash equilibrium for any* $\delta \in (0, 1]$*.*

Proof: see the appendix. □

As will be illustrated in the following example, both types of Nash equilibria described in Theorem 10 appear as the discount factor changes from 0 to 1. Moreover, for some range of discount factors, the Nash equilibrium is not unique.

Example 1. *Assume that* $b_1 = b_2 = 1$ *and* $f(x) = 1 - x^2$ *for* $x \in [0, 1]$*. Assume that the bargaining game is the simultaneous-offer game with symmetric arbitration.*

Analysis of the example: Table 2 lists the equilibrium type of the game and Figure 21 depicts the equilibrium payoff(s) of player 1. When $0 < \delta \leq 0.741$, there exist multiple Nash equilibria which include both the equilibrium with immediate agreement and the equilibrium with arbitration. Notice that although the equilibrium with arbitration is unique, the equilibrium with immediate agreement is not unique (except at $\delta = 0.741$). The range of player 1's payoffs obtained from equilibria with immediate agreement expands as the discount factor becomes small. As δ approaches zero, this range approaches $[0, 1]$, which means that any point on the Pareto frontier can be supported as the payoff of an immediate-agreement equilibrium.

δ	**Equilibrium Type**
$0 < \delta \leq 0.741$	Immediate-agreement, Arbitration
$0.741 < \delta \leq 1$	Arbitration

Table 2: Nash equilibrium of the game in Example 1.

When $0.741 < \delta \leq 1$, the unique Nash equilibrium is an equilibrium with arbitration. Notice that as δ approaches 1, the equilibrium payoff of player 1 converges to the payoff that he would receive from the Kalai-Smorodinsky (KS) solution outcome (we will further illustrate this point in Theorem 11). $\square$

In Example 1, when the discount factor is large, the unique Nash equilibrium is an equilibrium with arbitration; when the discount factor is small, then besides

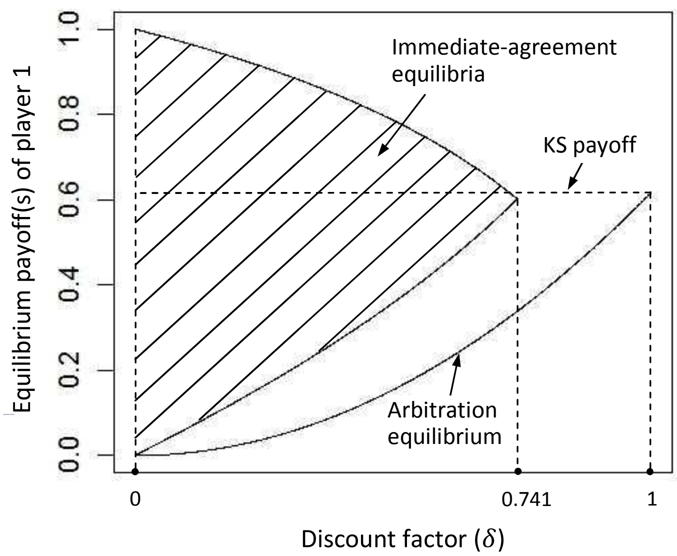

Figure 21: Equilibrium payoff(s) of player 1.

the equilibrium with arbitration, the equilibrium with immediate agreement also appears. These two results turn out to be general properties that are true for any bargaining set $S \in \mathcal{B}$. The next theorem (Theorem 11) summarizes these results.

Define $x^*(\delta)$ as the unique $x \in [0, b_1]$ that satisfies $\delta\tilde{x}(x) = x$, and $y^*(\delta)$ as the unique $y \in [0, b_2]$ that satisfies $\delta\tilde{y}(y) = y$. We have:

Theorem 11. *In the simultaneous-offer game with symmetric arbitration, there exists a $\hat{\delta} \in (0, 1)$, such that (i) if $\delta \in (0, \hat{\delta}]$, then for any $x \in [x^*(\delta), f^{-1}(y^*(\delta))]$ (which is nonempty), $((x, f(x)), (x, f(x)))$ is a Nash equilibrium;*[39] *and (ii) if $\delta \in (\hat{\delta}, 1]$, then $((b_1, 0), (0, b_2))$ is the only Nash equilibrium, and the stage 2 arbitrated outcome for the equilibrium $((b_1, 0), (0, b_2))$, $\gamma((b_1, 0), (0, b_2))$, coincides with the Kalai-Smorodinsky solution outcome of the Nash bargaining problem $((0, 0), S)$.*

Proof: see the appendix. □

[39] In addition, notice that $((b_1, 0), (0, b_2))$ is always a Nash equilibrium.

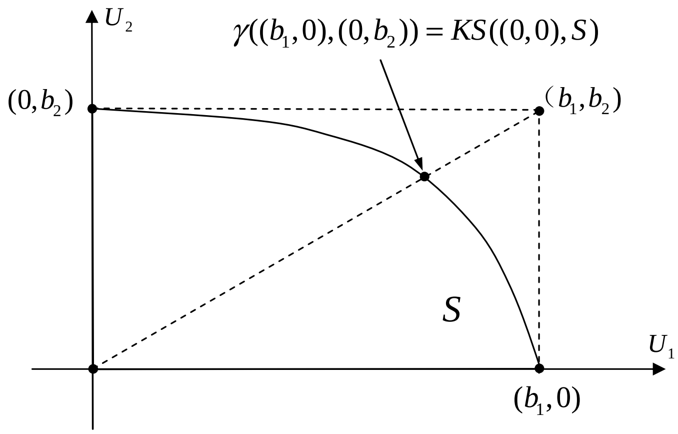

Figure 22: Equilibrium outcome when δ is close to 1.

According to Theorem 10 (ii), when the discount factor becomes close to 1, the unique equilibrium outcome of the simultaneous-offer game with symmetric arbitration coincides with the Kalai-Smorodinsky solution outcome of the Nash bargaining problem $((0,0),S)$. To see this, notice that the Kalai-Smorodinsky solution to the Nash bargaining problem $((0,0),S)$ is the intersection point of $L((0,0),(b_1,b_2))$ and the Pareto frontier (Kalai and Smorodinsky, 1975). The unique equilibrium outcome of our simultaneous-offer game when the discount factor is close to 1 is $\gamma((b_1,0),(0,b_2))$. Refer to Figure 22. It is obvious that $\gamma((b_1,0),(0,b_2)) = KS((0,0),S)$.

The key axiom that leads to the *symmetric arbitration solution* is the axiom of Symmetry in Offers and the key axiom that leads to the *Kalai-Smorodinsky solution* is the axiom of Individual Monotonicity.[40] Those two axioms have totally different

[40]The Kalai-Smorodinsky solution is the axiomatic solution that satisfies the following four axioms: Invariance w.r.t Affine Transformation, Pareto Optimality, Symmetry and Individual Monotonicity. The Kalai-Smorodinsky solution differs from the Nash solution by replacing the axiom of Independence of Irrelevant Alternatives with the axiom of Individual Monotonicity.

meanings and it is surprising that if we introduce arbitration in the simultaneous-offer game and require the arbitrator to obey the axiom of Symmetry in Offers (and the other two axioms), then the equilibrium outcome of the simultaneous-offer game will be the same as the Kalai-Smorodinsky solution outcome (as soon as δ is close to 1).

Corollary 3 of Anbarci and Boyd (2011) shows that when the continuation probability is *small*, the Kalai-Smorodinsky solution outcome must be one of the equilibrium outcomes. Moreover, the underlying equilibrium is an equilibrium with immediate agreement. Our result shows that when the discount factor is *large*, the Kalai-Smorodinsky solution outcome is the unique equilibrium outcome. Moreover, the underlying equilibrium is an equilibrium with arbitration.

When the bargaining set has a linear Pareto frontier, it can be verified that the threshold discount factor $\hat{\delta}$ in Theorem 11 is $\frac{2}{3}$, regardless of what the slope of the Pareto frontier might be. This threshold is the same as the threshold continuation probability obtained in Anbarci and Boyd (2011).[41] This is not surprising because (i) the continuation probability in Anbarci and Boyd (2011) is equivalent to the discount factor in our game (see also footnote 31), and (ii) the symmetric arbitration solution coincides with the rule of "equally splitting the difference" when the Pareto frontier is linear.

[41] The proof of corollary 3 of Anbarci and Boyd (2011) suggests that when the Pareto frontier is linear, the equilibrium with immediate agreement appears only if the continuation probability is less than $\frac{2}{3}$.

2.5.2 Alternating-Offer Game

This subsection will propose and analyze an alternating-offer game that involves symmetric arbitration.[42] In particular, we define the *alternating-offer game with symmetric arbitration* as the following three-stage procedure:

1. Stage 1: player 1 makes an offer $(x_1, y_1) \in PF(S)$ and player 2 decides whether to accept the offer, ending the game with (x_1, y_1), or reject the offer, moving the game on to the next stage;

2. Stage 2: player 2 makes an offer $(x_2, y_2) \in PF(S)$ and player 1 decides whether to accept the offer, ending the game with (x_2, y_2), or reject the offer, moving the game on to the final stage (arbitration stage);

3. Stage 3: an arbitrator decides the final outcome using the *symmetric arbitration solution*, i.e., $\gamma((x_1, y_1), (x_2, y_2))$ is the arbitrated outcome.[43]

Players' payoffs obtained at stage i is subject to a discount of δ^{i-1}. We will characterize the subgame perfect equilibria (henceforth SPE) of this game. We first impose two tie-breaking rules and make some definitions.

[42]Our game defined below is a variant of the alternating-offer game proposed by Yildiz (2011). Yildiz (2011) assumed that two players make offers sequentially and that if both offers are rejected by opponents, then the final offer arbitration rule is used to decide the final outcome. The final offer arbitration rule used by Yildiz (2011) is such that the offer that yields the higher Nash product is chosen as the arbitration outcome. It turns out that the unique subgame perfect equilibrium outcome in his game coincides with the equilibrium outcome in Rubinstein's infinite-horizon alternating-offer bargaining game (Rubinstein, 1982).

[43]We assume that if (x_1, y_1) and (x_2, y_2) are compatible, then each player gets what he demands at stage 2. Notice that in equilibrium, player 2 will never make an offer that is incompatible with player 1's offer.

Tie-breaking rule 1: whenever a player is indifferent between acceptance and rejection, he always chooses acceptance.

Tie-breaking rule 2: whenever a player is indifferent between the two options that he will offer his opponent, he always chooses the option that brings a higher payoff for his opponent.

Definition 12. *For any* $(x_1, y_1) \in PF(S)$ *with* $(x_1, y_1) \neq (0, b_2)$ *and* $(x_2, y_2) \in PF(S)$ *with* $x_2 \leq \delta x_1$*, define the following points (refer to Figure 23):* $A = (x_2, y_2)$*;* $B = (x_2, f(\frac{1}{\delta}x_2))$*;* $C = (\frac{1}{\delta}x_2, f(\frac{1}{\delta}x_2))$*;* $D = (\frac{1}{\delta}x_2, y_1)$ *and* $E = (x_1, y_1)$*.*

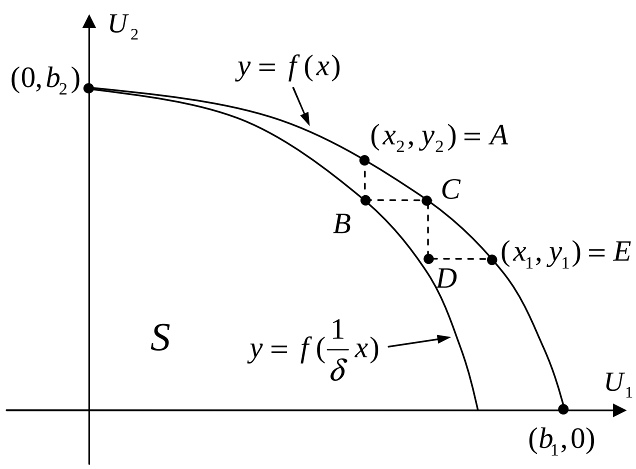

Figure 23: Definitions of points A, B, C, D and E.

The points A, B, C, D and E implicitly depend on (x_1, y_1) and/or (x_2, y_2). However, for simplicity, we omit that dependence in the notation. The curve $y = f(\frac{1}{\delta}x)$ in Figure 23 is obtained by fixing the payoff of player 2 and scaling down the payoff of player 1 by the discount factor δ.[44] Thus, for player 1, he must be

[44]To see this, note $y = f(\frac{1}{\delta}x)$ can be rewritten as $x = \delta f^{-1}(y)$.

indifferent between accepting the outcome B at the current stage and accepting the outcome C at the next stage. It should be noted that the point D is typically not on the curve $y = f(\frac{1}{\delta}x)$.

Definition 13. *For any given* $(x_1, y_1) \in PF(S)$ *with* $(x_1, y_1) \neq (0, b_2)$*, define* $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ *as the unique point* $(x_2, y_2) \in PF(S)$ *that satisfies: (i)* $x_2 \leq \delta x_1$*; (ii)* $|AB| * |BC| = |CD| * |DE|$.

The point $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ is well-defined because as $(x_2, y_2) \in PF(S)$ moves along the Pareto frontier from $(0, b_2)$ to $(\delta x_1, f(\delta x_1))$, $|AB| * |BC|$ strictly increases from zero to some positive number and $|CD| * |DE|$ strictly decreases from a positive number to zero. If $|AB| * |BC| = |CD| * |DE|$, then the main diagonal of the rectangle $AMEN$ must intersect the Pareto frontier at the point C (see Figure 24). That is, we must have $C = \gamma((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1)))$. Since for player 1, $\delta C \sim B$, we thus have $\delta\gamma_1((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))) = \hat{x}_2(x_1, y_1)$. The following lemma further shows that for any given (x_1, y_1), the point $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ is actually the only point on the Pareto frontier that satisfies $\delta\gamma_1((x_1, y_1), (x_2, y_2)) = x_2$. It also shows that $\hat{x}_2(x_1, y_1)$ is strictly increasing in x_1.

Lemma 14. *For* $(x_1, y_1) \in PF(S)$ *with* $(x_1, y_1) \neq (0, b_2)$*, we have:*

(i) $\delta\gamma_1((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))) = \hat{x}_2(x_1, y_1)$*;*

(ii) for any $(x_2, y_2) \in PF(S)$ *with* $x_2 < \hat{x}_2(x_1, y_1)$*, we have:* $\delta\gamma_1((x_1, y_1), (x_2, y_2)) > x_2$*;*

(iii) for any $(x_2, y_2) \in PF(S)$ *with* $x_2 > \hat{x}_2(x_1, y_1)$*, we have:* $\delta\gamma_1((x_1, y_1), (x_2, y_2)) < x_2$*;*

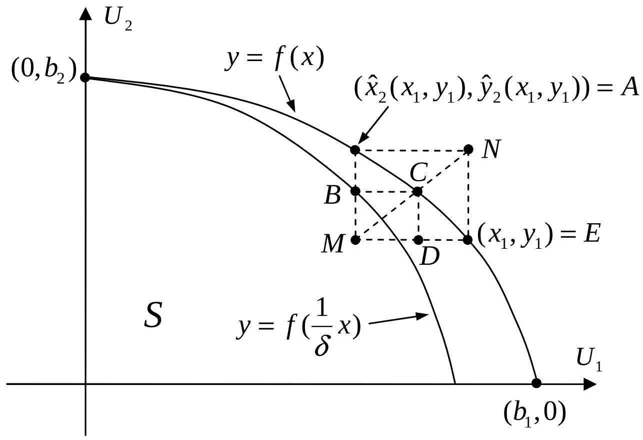

Figure 24: Definition of $(\hat{x}_2(x_1,y_1),\hat{y}_2(x_1,y_1))$.

(iv) if $(x_1',y_1')\in PF(S)$ *and* $x_1'>x_1$*, then we have:* $\hat{x}_2(x_1',y_1')>\hat{x}_2(x_1,y_1)$.

Proof: see the appendix. □

An implication of Lemma 14 (i) (ii) and (iii) is that, if the game were at stage 2 and player 2 made the offer (x_2,y_2), then whether or not player 1 accepts the offer depends on whether or not x_2 is greater than $\hat{x}_2(x_1,y_1)$. That is, we have:

Corollary 15. *Suppose player 1 offers* $(x_1,y_1)\neq(0,b_2)$ *at stage 1 which player 2 rejects and player 2 makes an offer* (x_2,y_2) *at stage 2, then player 1 will accept the offer* (x_2,y_2) *if and only if* $x_2\geq\hat{x}_2(x_1,y_1)$.

The following lemma characterizes the players' equilibrium behavior at stage 2. It is essential for our main result in characterizing the SPE of our entire game.

Lemma 16. *In equilibrium, if at stage 1, player 1 offers* $(x_1,y_1)\neq(0,b_2)$ *which player 2 rejects, then at stage 2, we have:*

(i) player 2 must either offer $(0, b_2)$ *which player 1 rejects, or offer* $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ *which player 1 accepts.*

(ii) if $(x_1, y_1) \neq (b_1, 0)$*, then player 2 must be indifferent between offering* $(0, b_2)$ *and offering* $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$*, i.e.,* $\delta\gamma_2((x_1, y_1), (0, b_2)) = \hat{y}_2(x_1, y_1)$*.*

Proof: see the appendix. □

The intuition of Lemma 16 (i) is straightforward. If player 1 offers $(x_1, y_1) \neq (0, b_2)$ which player 2 rejects, then at stage 2, player 2 can either make an offer that player 1 will reject or make an offer that player 1 will accept. In the former case, player 2's best option is to make the extreme offer $(0, b_2)$, because the more demand he makes in his offer, the more arbitrated payoff he can obtain at the arbitration stage (according to Lemma 8). In the latter case, player 2's best option is to make the offer $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$, because his offer (x_2, y_2) will be accepted by player 1 if and only if $x_2 \geq \hat{x}_2(x_1, y_1)$ (according to Corollary 15).

Lemma 16 (ii) states that as soon as $(x_1, y_1) \notin \{(0, b_2), (b_1, 0)\}$ is rejected by player 2 at stage 1, then player 2 must be indifferent between making the extreme offer (i.e., offering $(0, b_2)$) and "concession" (i.e., offering $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$) at stage 2. This is because, if player 2 strictly prefers one option over the other, say, player 2 strictly prefers offering $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ over offering $(0, b_2)$, then at stage 1, player 1 has the incentive to deviate to a slightly more extreme offer $(x_1+\epsilon, f(x_1+\epsilon))$. Such a small deviation will not change player 2's preference over the two options at stage 2, i.e., player 2 strictly prefers offering $(\hat{x}_2(x_1+\epsilon, f(x_1+\epsilon)), \hat{y}_2(x_1+\epsilon, f(x_1+\epsilon)))$ over offering $(0, b_2)$. As a result, after deviation, player 1 obtains a payoff of $\hat{x}_2(x_1 + \epsilon, f(x_1 + \epsilon))$ which is higher than $\hat{x}_2(x_1, y_1)$, the payoff before deviation (according to Lemma 14 (iv)).

A direct result of Lemma 16 (i) is the following theorem, which characterizes all SPE of the game. Note that in equilibrium, player 1 will never offer $(0, b_2)$ at stage 1 because the offer $(0, b_2)$ is dominated by the offer $(b_1, 0)$ which will bring him a payoff of at least $\delta^2\gamma((b_1, 0), (0, b_2)) > 0$. In addition, using tie-breaking rule 1 and tie-breaking rule 2, it can be shown that the SPE of the game must be unique.

Theorem 17. *In the alternating-offer game with symmetric arbitration, there exists a unique SPE and the unique SPE must take one of the following three forms:*

(i) (immediate-agreement) at stage 1, player 1 offers $(x_1, y_1) \neq (0, b_2)$ which player 2 accepts;

(ii) (delayed-agreement) at stage 1, player 1 offers $(x_1, y_1) \neq (0, b_2)$ which player 2 rejects; at stage 2, player 2 offers $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ which player 1 accepts;

(iii) (arbitration) at stage 1, player 1 offers $(x_1, y_1) \neq (0, b_2)$ which player 2 rejects; at stage 2, player 2 offers $(0, b_2)$ which player 1 rejects.

Theorem 17 states that the unique SPE of the alternating-offer game with symmetric arbitration is either an equilibrium with immediate agreement, or an equilibrium with delayed agreement, or an equilibrium with arbitration. In the following example, all of the three forms of SPE appear when the discount factor changes from 0 to 1.

Example 2. *Assume that $b_1 = b_2 = 1$ and $f(x) = 1 - x^2$ for $x \in [0, 1]$. Assume that the bargaining game is the alternating-offer game with symmetric arbitration.*

δ	**Equilibrium Type**
$0 < \delta \leq 0.808$	Immediate-agreement
$0.808 < \delta < 0.834$	Delayed-agreement
$0.834 \leq \delta \leq 0.902$	Immediate-agreement
$0.902 < \delta \leq 1$	Arbitration

Table 3: SPE of the game in Example 2.

Analysis of the example: Table 3 lists the equilibrium type of the game. When the discount factor is small, the unique SPE is an equilibrium with immediate agreement. When the discount factor is large, the unique SPE is an equilibrium with arbitration. These two properties turn out to be two general properties that hold for any bargaining set in $S \in \mathcal{B}$.[45] This point is shown in Theorem 18. $\square$

For any given $\delta \in (0, 1]$, define $x_1^{**}(\delta)$ as the unique $x_1 \in [0, b_1]$ that satisfies $f(x_1) = \max\{\delta^2 \tilde{y}(f(x_1)), \delta \hat{y}_2(x_1, f(x_1))\}$.[46] We have the following theorem which characterizes the SPE of the alternating-offer game with symmetric arbitration for the cases where the discount factor is either sufficiently small or sufficiently large.

Theorem 18. *There exists a $\delta^{**} \in (0, 1)$ and a $\delta^* \in (0, 1)$ with $0 < \delta^{**} \leq \delta^* < 1$, such that (i) when $\delta \in (0, \delta^{**})$, the unique SPE of the alternating-offer game with symmetric arbitration is that at stage 1, player 1 makes the offer $(x_1^{**}(\delta), f(x_1^{**}(\delta)))$ which player 2 accepts immediately, and (ii) when $\delta \in (\delta^*, 1]$, the unique SPE of the*

[45]These results are similar to the results obtained by Rong (2011) for the alternating-offer game with equally-split-the-difference arbitration in the sense that in both games, when the discount factor is small, the equilibrium features immediate agreement, and when the discount factor is large, the equilibrium features arbitration. This is not surprising because both the symmetric arbitration and equally-split-the-difference arbitration have the common feature that they allow compromise between offers. However, the equilibrium outcomes for the two games are different.

[46]It can be verified that player 2 will accept player 1's offer $(x, f(x))$ if and only if $x \leq x^{**}(\delta)$.

alternating-offer game with symmetric arbitration is that at stage 1, player 1 makes the offer $(b_1, 0)$ *which player 2 rejects, and at stage 2, player 2 makes the offer* $(0, b_2)$ *which player 1 rejects; the equilibrium outcome of the game is thus* $\gamma((b_1, 0), (0, b_2))$ *which coincides with the Kalai-Smorodinsky solution outcome of the Nash bargaining problem* $((0, 0), S)$.

Proof: see the appendix. $\square$

When the Pareto frontier of the bargaining set is linear, it can be verified that the threshold discount factor δ^* is 0.91. This threshold is much larger than the threshold discount factor obtained in the simultaneous-offer game for the linear Pareto frontier case.[47] This is because there are three stages in the alternating-offer game, but only two stages in the simultaneous-offer game. In other words, for a given discount factor, the arbitration outcome is discounted more severely in the alternating-offer game. As a result, in the alternating-offer game, players have less incentive to make extreme offers and the result that players make extreme offers in equilibrium is less robust.

2.6 Conclusion

This paper defines a class of arbitration problems and characterizes its solution using the axiomatic approach. We impose three axioms that an arbitrator should use. They are "Symmetry in Offers", "Invariance" and "Pareto Optimality". The key rule, Symmetry in Offers, requires that whenever players' offers are symmetric,

[47]The corresponding threshold discount factor in Example 2 is also greater than that in Example 1.

the arbitrated outcome should also be symmetric. We show that there is a unique arbitration solution, called the symmetric arbitration solution, that satisfies all three axioms. The symmetric arbitration solution has a simple graphical representation.

We then introduce symmetric arbitration in two bargaining games. One is the simultaneous-offer game and the other is the alternating-offer game. At the arbitration stage of both games, the arbitrator uses the symmetric arbitration solution to decide the arbitration outcome. We show that in both games, if the discount factor is sufficiently small, the equilibrium with immediate agreement will appear. If the discount factor is sufficiently close to 1, then the *unique* equilibrium is such that both players make extreme offers and the corresponding equilibrium outcome is the Kalai-Smorodinsky solution outcome.

Although the equilibrium outcomes of the two games coincide with that of the Kalai-Smorodinsky solution (when δ is close to 1), our result is not a typical implementation result. Strictly speaking, a strategic implementation of an axiomatic bargaining solution requires that the mechanism used for implementation can be translated into a form that only depends on the physical outcomes of bargaining and not on the players' preferences or utility representations (Serrano, 1997; Dagan and Serrano, 1998). Our games cannot be translated into a form that only depends on the physical outcomes, because the symmetric arbitration solution is defined on the basis of players' utilities. However, compared with the implementation mechanism on the Kalai-Smorodinsky solution in the literature[48], our games are much more simple. Our games also help us to understand the Kalai-Smorodinsky solution from

[48]See, for example, the "auctioning fractions of dictatorship" mechanism proposed by Moulin (1984).

a new perspective. That is, the Kalai-Smorodinsky solution outcome is the only fair and efficient arbitration outcome when both players make extreme offers.

Although our model assumes that both players have the same discount factor, our main result can be extended to the case where the two players' discount factors differ. That is, as long as the discount factors of the two players are sufficiently close to 1, then the unique equilibrium outcome of both the simultaneous-offer game and the alternating-offer game coincides with the Kalai-Smorodinsky solution outcome.[49]

2.7 Chapter 2 Appendix

Proof of Lemma 8:

Let's consider γ_1. For any given $x_1 \in [d_1, b_1]$ and $x_1' \in [d_1, b_1]$ with $x_1 < x_1'$ and any $x_2 \in [d_2, b_2]$ with $x_2 < x_1$, the line connecting $(x_1, f(x_1)) \wedge (x_2, f(x_2))$ and $(x_1, f(x_1)) \vee (x_2, f(x_2))$ is strictly above the line connecting $(x_1', f(x_1')) \wedge (x_2, f(x_2))$ and $(x_1', f(x_1')) \vee (x_2, f(x_2))$ (see Figure 25). Since the Pareto frontier is strictly downward-sloping, we must have $\gamma_1((x_1, f(x_1)), (x_2, f(x_2))) < \gamma_1((x_1', f(x_1')), (x_2, f(x_2)))$. Thus, $\gamma_1((x_1, f(x_1)), (x_2, f(x_2)))$ is strictly increasing in x_1. Similarly, $\gamma_1((x_1, f(x_1)), (x_2, f(x_2)))$ is strictly increasing in x_2.

The proof for γ_2 is similar and is omitted.

Proof of Theorem 10:

(i) The result holds if the following is true:

[49] If both discount factors are small, then the equilibrium with immediate agreement appears. If the discount factor of one player is large and the discount factor of the other player is small, then the equilibrium type may depend on the discount factors and the shape of the bargaining set in a complex manner.

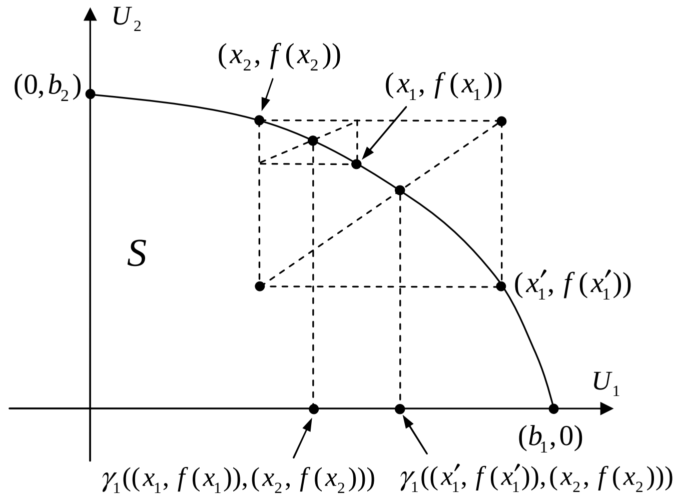

Figure 25: $\gamma_1((x_1, f(x_1)), (x_2, f(x_2)))$ and $\gamma_1((x_1', f(x_1')), (x_2, f(x_2)))$ where $x_1 < x_1'$.

(a) if player 2's offer is $(x, y) \in PF(S)$, then (x, y) is player 1's best response if and only if $x \geq \delta\tilde{x}(x)$; (b) if player 1's offer is $(x, y) \in PF(S)$, then (x, y) is player 2's best response if and only if $y \geq \delta\tilde{y}(y)$.

We will only prove (a) in the following. The proof of (b) is similar.

Suppose player 2's offer is (x, y). If player 1 makes an offer $(x', y') \in PF(S)$ with $0 \leq x' < x$, then since (x', y') and (x, y) are compatible, player 1's payoff must be x', which is strictly less than x. If player 1 makes an offer $(x', y') \in PF(S)$ with $x < x' \leq b_1$, then by Lemma 8, his payoff is at most $\delta\tilde{x}(x)$. Thus, we have shown that (x, y) is player 1's best response if and only if $x \geq \delta\tilde{x}(x)$.

(ii) We will show that for any $\delta \in (0, 1]$, $((b_1, 0), (0, b_2))$ is a Nash equilibrium. We will first show that $(b_1, 0)$ is player 1's best response to player 2's offer $(0, b_2)$. Suppose player 2's offer is $(0, b_2)$, then player 1 can either make the offer $(0, b_2)$ or make some offer $(x, f(x)) \neq (0, b_2)$. If player 1 offers $(0, b_2)$, then his payoff is 0. If player 1 offers $(x, f(x)) \neq (0, b_2)$, then the game will move to the arbitration stage and player 1's payoff is $\delta\gamma_1((x, f(x)), (0, b_2)) > 0$. Now, by Lemma 8, $\delta\gamma_1((x, y), (0, b_2))$ is strictly increasing in x, so player 1's best response to player 2's offer $(0, b_2)$ is $(b_1, 0)$. Similarly, player 2's best response to player 1's offer $(b_1, 0)$ is $(0, b_2)$. Thus, $((b_1, 0), (0, b_2))$ is a Nash equilibrium for any $\delta \in (0, 1]$.

At last, note that $((x, y), (x, y))$ $((x, y) \in PF(S))$ and $((b_1, 0), (0, b_2))$ are the only two possible types of Nash equilibria, i.e., any $((x_1, f(x_1)), (x_2, f(x_2)))$ with $((x_1, f(x_1)), (x_2, f(x_2))) \neq ((b_1, 0), (0, b_2))$ and $x_1 \neq x_2$ cannot be the Nash equilibrium for any $\delta \in (0, 1]$. This is because (i) if $x_1 < x_2$, then the two offers are compatible and player 1 has incentive to deviate to $(x_1 + \epsilon, f(x_1 + \epsilon))$ with some $x_1 + \epsilon \leq x_2$; (ii) if $x_1 > x_2$, then by Lemma 8, the player who does not make the extreme offer has the incentive to deviate to making the extreme offer.

Proof of Theorem 11:

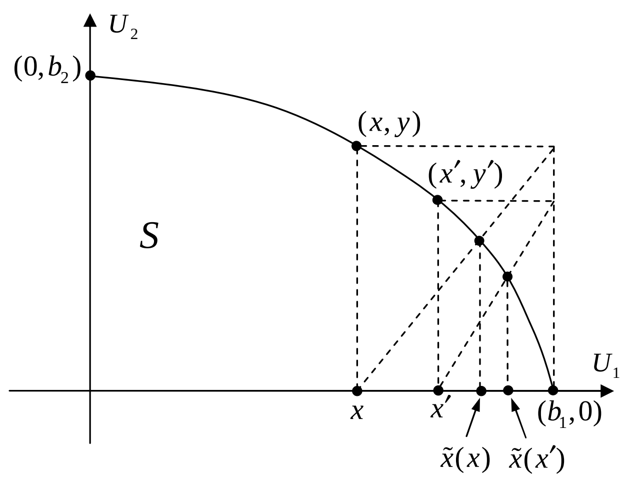

Figure 26: $\tilde{x}(x) - x$ is strictly decreasing in x and $\tilde{x}(x)$ is strictly increasing in x.

First notice that $x^*(\delta)$ is well-defined because (i) $\delta\tilde{x}(x) - x = (\tilde{x}(x) - x) - (1 - \delta)\tilde{x}(x)$ is strictly decreasing in x (see Figure 26), (ii) $\delta\tilde{x}(x) - x > 0$ at $x = 0$, and

(iii) $\delta\tilde{x}(x) - x \leq 0$ at $x = b_1$. Similarly, $y^*(\delta)$ is well-defined.

Since $\delta\tilde{x}(x) - x$ is strictly decreasing in x, $x \geq \delta\tilde{x}(x)$ if and only if $x \geq x^*(\delta)$. That is, if player 2's offer is $(x, y) \in PF(S)$, then player 1's best response is to make the same offer if and only if $x \geq x^*(\delta)$. Similarly, if player 1's offer is $(x, y) \in PF(S)$, then player 2's best response is to make the same offer if and only if $y \geq y^*(\delta)$ (see Figure 27).

Observing that $x^*(\delta) \to b_1$ and $f^{-1}(y^*(\delta)) \to 0$ as $\delta \to 1$, and $x^*(\delta) \to 0$ and $f^{-1}(y^*(\delta)) \to b_1$ as $\delta \to 0$, there exists a unique $\delta \in (0, 1)$, denoted by $\hat{\delta}$, that satisfies $x^*(\delta) = f^{-1}(y^*(\delta))$.

According to Lemma 10, $((x, y), (x, y))$ $((x, y) \in PF(S))$ is a Nash equilibrium if and only if $x \geq \delta\tilde{x}(x)$ and $y \geq \delta\tilde{y}(y)$. So, $((x, y), (x, y))$ $((x, y) \in PF(S))$ is a Nash equilibrium if and only if $x^*(\delta) \leq x \leq f^{-1}(y^*(\delta))$. The remainder of the proof is straightforward and is omitted. $\square$

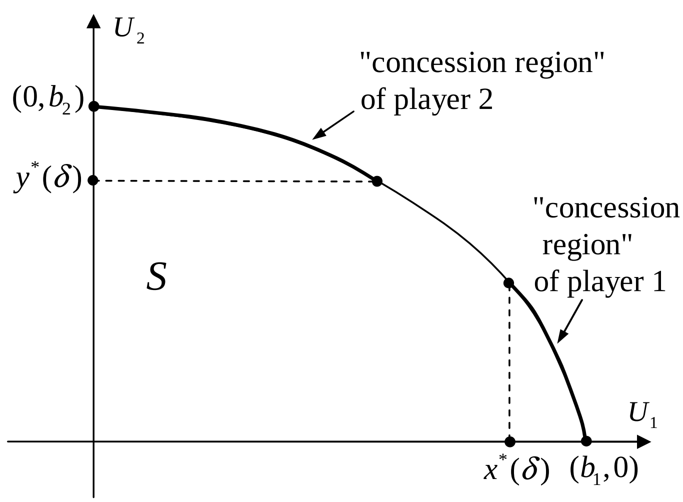

Figure 27: The regions that players will choose "concession" instead of making the extreme offers.

Proof of Lemma 14:

(i) Refer to Figure 24. By definition, for any given (x_1, y_1) with $(x_1, y_1) \neq (0, b_2)$, the pair $((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1)))$ satisfies $|AB| * |BC| = |CD| * |DE|$. This implies that the point C must be on the line that connects $(x_1, y_1) \wedge (x_2, y_2)$ (i.e., M) and $(x_1, y_1) \vee (x_2, y_2)$ (i.e., N). In addition, notice that point C is on $PF(S)$. Then, we must have $C = \gamma((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1)))$. Since by definition, $C = (\frac{1}{\delta}\hat{x}_2(x_1, y_1), f(\frac{1}{\delta}\hat{x}_2(x_1, y_1)))$, we have:

$$\delta\gamma_1((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))) = \hat{x}_2(x_1, y_1).$$

(ii) Refer to Figure 28. Note that for given (x_1, y_1) with $(x_1, y_1) \neq (0, b_2)$, as (x_2, y_2) moves from the lower-right to the upper-left along the Pareto frontier, the corresponding $|AB| * |BC|$ strictly decreases and $|CD| * |DE|$ strictly increases. Thus, for $(x_2, y_2) \in PF(S)$ with $x_2 < \hat{x}_2(x_1, y_1)$, we must have $|AB| * |BC| < |CD| * |DE|$. This implies the slope of the line MC is bigger than that of the line MN. Thus, the point O (the intersection point of the line MN with $PF(S)$) must be on the right of the line CD, then we have: $\gamma_1((x_1, y_1), (x_2, y_2)) > \frac{1}{\delta}x_2$, i.e., $\delta\gamma_1((x_1, y_1), (x_2, y_2)) > x_2$.

(iii) We have three sub-cases here:

(a) $\hat{x}_2(x_1, y_1) < x_2 \leq \delta x_1$

Refer to Figure 29. Note that for any given (x_1, y_1) with $(x_1, y_1) \neq (0, b_2)$, as (x_2, y_2) moves from the upper-left to the lower-right along the Pareto frontier, the corresponding $|AB| * |BC|$ strictly increases and $|CD| * |DE|$ strictly decreases. Thus, for $(x_2, y_2) \in PF(S)$ with $x_2 > \hat{x}_2(x_1, y_1)$, we must have $|AB| * |BC| > |CD| * |DE|$.

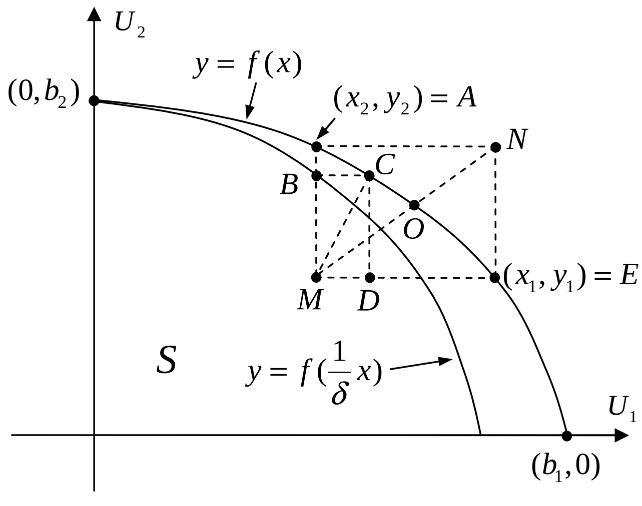

Figure 28: The case where $x_2 < \hat{x}(x_1, y_1)$.

This implies the slope of the line MC is smaller than that of the line MN. Thus, the point O must be on the left of the line CD, then we have: $\gamma_1((x_1, y_1), (x_2, y_2)) < \frac{1}{\delta}x_2$, i.e., $\delta\gamma_1((x_1, y_1), (x_2, y_2)) < x_2$.

(b) $\delta x_1 < x_2 \leq x_1$

For this case, since $x_2 \leq x_1$, then we must have $\gamma_1((x_1, y_1), (x_2, y_2)) \leq x_1$. Then: $\delta\gamma_1((x_1, y_1), (x_2, y_2)) \leq \delta x_1 < x_2$.

(c) $x_2 > x_1$

For this case, since $x_2 > x_1$, then we have $\gamma_1((x_1, y_1), (x_2, y_2)) < x_2$. Thus, $\delta\gamma_1((x_1, y_1), (x_2, y_2)) < x_2$.

(iv) Refer to Figure 30. Suppose we have $(x_1', y_1') \in PF(S)$ and $x_1' > x_1$. Now, for (x_1, y_1) and $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$, we have:

$$|AB| * |BC| = |CD| * |DE|$$

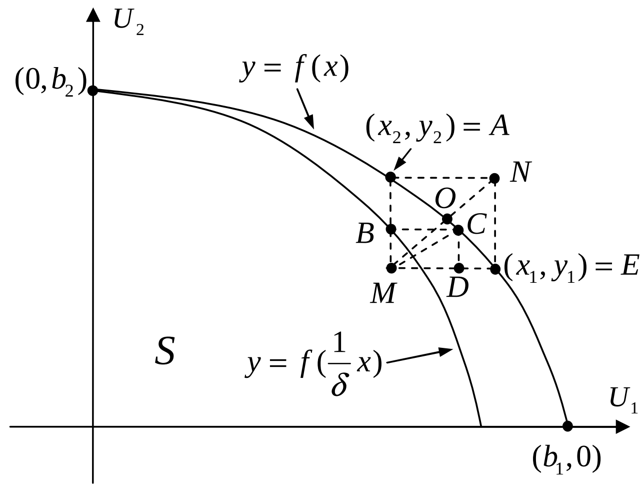

Figure 29: The case where $\hat{x}_2(x_1, y_1) < x_2 \leq \delta x_1$.

Since (x_1', y_1') is on the lower right of (x, y), we have:

$$|AB| * |BC| < |CD'| * |D'E'|.$$

Again, note that for given (x_1, y_1), as (x_2, y_2) moves from the upper-left to the lower-right along the Pareto frontier, $|AB| * |BC|$ strictly increases, and $|BC| * |CD|$ strictly decreases. So, we must have

$$\hat{x}_2(x_1', y_1') > \hat{x}_2(x_1, y_1).$$

Proof of Lemma 16:

(i) The proof is obvious and is omitted.

(ii) Suppose player 1 offers $(x_1, y_1) \notin \{(0, b_2), (b_1, 0)\}$. By part (i), we know that if player 2 rejects $(x_1, y_1) \neq (0, b_2)$, then at stage 2, he must either offer $(0, b_2)$ or

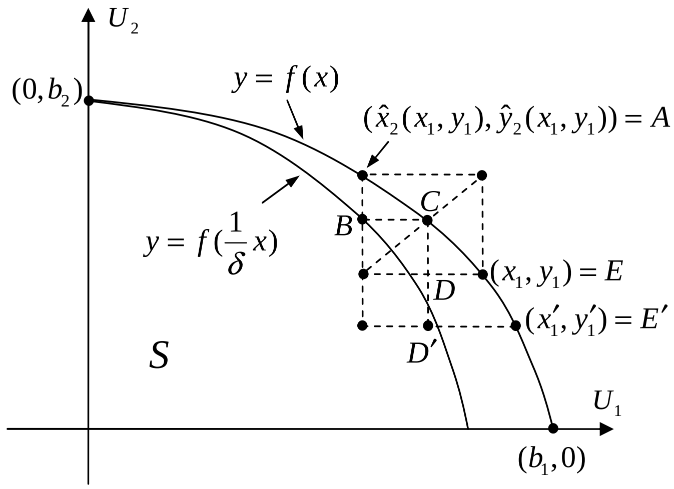

Figure 30: A comparison of $\hat{x}_2(x_1', y_1')$ and $\hat{x}_2(x_1, y_1)$ where $x_1' > x_1$.

offer $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$. The corresponding (stage 1) payoff for player 2 is either $\delta^2\gamma_2((x_1, y_1), (0, b_2))$ or $\delta\hat{y}_2(x_1, y_1)$. Since player 2 chooses to reject (x_1, y_1) at stage 1, then we must have:

$$max\{\delta^2\gamma_2((x_1, y_1), (0, b_2)), \delta\hat{y}_2(x_1, y_1)\} > y_1.$$

Note that the above inequality is strict because we have assumed that whenever a player is indifferent between "accept" and "reject", he must choose "accept". Now, let's consider the following two cases:

(a) $\delta^2\gamma_2((x_1, y_1), (0, b_2)) > \delta\hat{y}_2(x_1, y_1)$

In this case, player 2 must offer $(0, b_2)$ at stage 2. Player 1 thus obtains a payoff of $\delta^2\gamma_1((x_1, y_1), (0, b_2))$. We will show that player 1 will gain more if he submits a more extreme offer at stage 1. In particular, since $(x_1, y_1) \neq (b_1, 0)$, we can find an $\epsilon' > 0$ such that $x_1' = x_1 + \epsilon' < b_1$, $y_1' = f(x_1')$, $\delta^2\gamma_2((x_1', y_1'), (0, b_2)) > \delta\hat{y}_2(x_1', y_1')$

and $max\{\delta^2\gamma_2((x_1', y_1'), (0, b_2)), \delta\hat{y}_2(x_1', y_1')\} > y_1'$. That is, if player 1 offers (x_1', y_1') at stage 1, then player 2 must reject it and still offer $(0, b_2)$ at stage 2. Thus, player 1 will obtain a payoff of $\delta^2\gamma_1((x_1', y_1'), (0, b_2))$ by offering (x_1', y_1') at stage 1. Now, since $x_1' > x_1$, we have $\delta^2\gamma_1((x_1', y_1'), (0, b_2)) > \delta^2\gamma_1((x_1, y_1), (0, b_2))$. That is, player 1 is better off by offering (x_1', y_1').

(b) $\delta\hat{y}_2(x_1, y_1) > \delta^2\gamma_2((x_1, y_1), (0, b_2))$

In this case, player 2 must offer $(\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))$ at stage 2. Player 1 obtains a payoff of $\delta\hat{x}_2(x_1, y_1)$. Again, we will show that player 1 can gain more if he makes a more extreme offer at stage 1. In particular, since $(x_1, y_1) \neq (b_1, 0)$, we can find an $\epsilon'' > 0$ such that $x_1'' = x_1 + \epsilon'' < b_1$, $y_1'' = f(x_1'')$, $\delta\hat{y}_2(x_1'', y_1'') > \delta^2\gamma_2((x_1'', y_1''), (0, b_2))$ and $max\{\delta^2\gamma_2((x_1'', y_1''), (0, b_2)), \delta\hat{y}_2(x_1'', y_1'')\} > y_1''$. That is, if player 1 offers (x_1'', y_1'') at stage 1, then player 2 must reject it and offer $(\hat{x}_2(x_1'', y_1''), \hat{y}_2(x_1'', y_1''))$ at stage 2. Note $(\hat{x}_2(x_1'', y_1''), \hat{y}_2(x_1'', y_1''))$ must be accepted by player 1. Thus, player 1 will obtain a payoff of $\delta\hat{x}_2(x_1'', y_1'')$ by offering (x_1'', y_1'') at stage 1. Now, by Lemma 14 (iv), since $x_1'' > x_1$, we have: $\delta\hat{x}_2(x_1'', y_1'') > \delta\hat{x}_2(x_1, y_1)$. That is, player 1 is better off by offering (x_1'', y_1'').

Thus, we have proved that if $(x_1, y_1) \notin \{(0, b_2), (b_1, 0)\}$, then we must have $\delta\gamma_2((x_1, y_1), (0, b_2)) = \hat{y}_2(x_1, y_1)$ in equilibrium.

Proof of Theorem 18:

In the following, we will prove part (ii). The idea of proof for part (i) is similar and is thus omitted.

We denote $\gamma((b_1, 0), (0, b_2))$ as (x^*, y^*) for simplicity.

Let δ_1^* be the unique $\delta \in (0, 1)$ that satisfies $\delta^2 = \dfrac{2f(\delta^2 x^*)}{f(\delta^2 x^*) + b_2}$. Let $\delta_2^* =$

$\max_{\frac{4}{9}x^* \leq x_1 \leq b_1} \delta_2^*(x_1)$ where $\delta_2^*(x_1)$ is the unique $\delta \in (0,1)$ that satisfies $\delta \dfrac{f(x_1)+b_2}{2} = f(\dfrac{\frac{\delta}{2}x_1}{1-\frac{\delta}{2}})$. Let $\delta^* = max\{\delta_1^*, \delta_2^*, \frac{2}{3}\}$. Note since $\delta_1^* \in (0,1)$ and $\delta_2^* \in (0,1)$, we have $\delta^* \in (0,1)$.

The proof is divided into two steps.

First step: We will show that, if $\delta \in (\delta^*, 1]$, then player 1 must offer $(b_1, 0)$ at stage 1.

First, we will show that, if $\delta \in (\delta^*, 1]$, then player 1 will never offer (x_1, y_1) with $x_1 \in [0, \delta^2 x^*)$. In particular, We will show that for player 1, any offer (x_1, y_1) with $x_1 \in [0, \delta^2 x^*)$ is strictly dominated by the offer $(b_1, 0)$.

Note that if player 1 makes the offer (x_1, y_1) with $x_1 \in [0, \delta^2 x^*)$, then his payoff is at most x_1. If player 1 proposes $(b_1, 0)$, then by Lemma 16, player 2 may choose to accept, or reject with counteroffer $(0, b_2)$ which player 1 rejects, or reject with counteroffer $(\hat{x}_2(b_1,0), \hat{y}_2(b_1,0))$ which player 1 accepts. If player 2 accepts, then player 1's payoff is b_1; if player 2 rejects with counteroffer $(0, b_2)$ which player 1 rejects, then player 1's payoff is $\delta^2\gamma_1((b_1,0),(0,b_2)) = \delta^2 x^*$; if player 2 rejects with counteroffer $(\hat{x}_2(b_1,0), \hat{y}_2(b_1,0))$ which player 1 will accept, then player 1's payoff is $\delta\hat{x}_2(b_1,0) = \delta^2\gamma_1((b_1,0),(\hat{x}_2(b_1,0),\hat{y}_2(b_1,0)))$ (the equality is by Lemma 14 (i)). Since $\gamma_1((b_1,0),(\hat{x}_2(b_1,0),\hat{y}_2(b_1,0))) \geq \gamma_1((b_1,0),(0,b_2))$, we have: $\delta\hat{x}_2(b_1,0) \geq \delta^2\gamma_1((b_1,0),(0,b_2)) = \delta^2 x^*$. Thus, we have shown that, if player 1 proposes $(b_1, 0)$ at stage 1, then his payoff is at least $\delta^2 x^*$. Hence, player 1 will never offer (x_1, y_1) with $x_1 \in [0, \delta^2 x^*)$.

Second, we will show that, if $\delta \in (\delta^*, 1]$, then player 1 will never offer (x_1, y_1) with $x_1 \in [\delta^2 x^*, b_1)$. We have the following two cases:

(i) (x_1, y_1) is accepted by player 2.

Note that, if player 2 rejects (x_1, y_1), then his payoff is at least $\delta^2\gamma_2((x_1, y_1), (0, b_2))$. Since player 2 chooses to accept (x_1, y_1), then we must have:

$$\delta^2\gamma_2((x_1, y_1), (0, b_2)) \leq y_1.$$

Since $\delta^2\gamma_2((x_1, y_1), (0, b_2)) \geq \frac{\delta^2}{2}(y_1 + b_2)$ (using the fact that the Pareto frontier is strictly "bowed-out"), then we have $\frac{\delta^2}{2}(y_1 + b_2) \leq y_1$, i.e., $\delta^2 \leq \dfrac{2y_1}{y_1 + b_2}$. Since $y_1 = f(x_1) \leq f(\delta^2 x^*) \leq f(\delta_1^{*2} x^*)$ and $\dfrac{2y_1}{y_1 + b_2}$ is increasing in y_1, we have $\delta^2 \leq \dfrac{2f(\delta_1^{*2}x^*)}{f(\delta_1^{*2}x^*) + b_2} = \delta_1^{*2} \leq \delta^{*2}$. Contradiction with $\delta > \delta^*$.

(ii) (x_1, y_1) is rejected by player 2.

We will compare $\delta\gamma_1((x_1, y_1), (0, b_2))$ and $\hat{y}_2(x_1, y_1)$.

First, note that by Lemma 14 (i) and the fact that the Pareto frontier is strictly "bowed-out", we have: $\hat{x}_2(x_1, y_1) = \delta\gamma_1((x_1, y_1), (\hat{x}_2(x_1, y_1), \hat{y}_2(x_1, y_1))) \geq \delta\dfrac{\hat{x}_2(x_1, y_1) + x_1}{2}$. Then we have: $\hat{x}_2(x_1, y_1) \geq \dfrac{\frac{\delta}{2}x_1}{1 - \frac{\delta}{2}}$. Then, $\hat{y}_2(x_1, y_1) \leq f(\dfrac{\frac{\delta}{2}x_1}{1 - \frac{\delta}{2}})$. Therefore,

$$\hat{y}_2(x_1, y_1) \leq f(\frac{\frac{\delta}{2}x_1}{1 - \frac{\delta}{2}}) < \delta\frac{y_1 + b_2}{2} \tag{8}$$

The last inequality is because $\delta > \delta^* \geq \delta_2^* \geq \delta_2^*(x_1)$,[50] $\delta_2^*(x_1)$ satisfies $f(\dfrac{\frac{\delta_2^*(x_1)}{2}x_1}{1 - \frac{\delta_2^*(x_1)}{2}}) = \delta_2^*(x_1)\dfrac{f(x_1) + b_2}{2}$, $f(\dfrac{\frac{\delta}{2}x_1}{1 - \frac{\delta}{2}})$ is strictly decreasing in δ, $\delta\dfrac{y_1 + b_2}{2}$ is strictly increasing in δ and $y_1 = f(x_1)$.

[50] The inequality $\delta_2^* \geq \delta_2^*(x_1)$ is true because of the following. Note that $\delta > \delta^*$ implies $\delta > 2/3$, which implies $\delta_2^* = \max_{\frac{4}{9}x^* \leq x_1 \leq b_1} \delta_2^*(x_1) \geq \max_{\delta^2 x^* \leq x_1 \leq b_1} \delta_2^*(x_1)$. So, for any $x_1 \in [\delta^2 x^*, b_1)$, we have $\delta_2^* \geq \delta_2^*(x_1)$.

Now, note $\delta\gamma_2((x_1, y_1), (0, b_2)) \geq \delta\dfrac{y_1 + b_2}{2}$. Then, we have:

$$\delta\gamma_2((x_1, y_1), (0, b_2)) > \hat{y}_2(x_1, y_1).$$

However, by Lemma 16 (ii), if at stage 1, player 1's makes the offer $(x_1, y_1) \notin \{(0, b_2), (b_1, 0)\}$ and player 2 rejects it, then we must have

$$\delta\gamma_2((x_1, y_1), (0, b_2)) = \hat{y}_2(x_1, y_1).$$

Contradiction!

Thus, we have proved that player 1 will never offer (x_1, y_1) with $x_1 \in [\delta x^*, 1)$. In addition, we have already proved that player 1 will never offer (x_1, y_1) with $x_1 \in [0, \delta x^*)$. Thus, player 1 must offer $(b_1, 0)$ at Stage 1.

Second Step: We will show that if $\delta \in (\delta^*, 1]$, and if player 1 offers $(b_1, 0)$ at stage 1, then player 1 must reject it and offers $(0, b_2)$ at stage 2.

If $(b_1, 0)$ is proposed by player 1 at stage 1, then by Lemma 14, player 2 has three options: (a) accept $(b_1, 0)$ – player 2's payoff is 0; (b) reject $(b_1, 0)$ and makes the counteroffer $(0, b_2)$ – player 2's payoff is $\delta^2\gamma_2((b_1, 0), (0, b_2))$; and (c) reject $(b_1, 0)$ and makes the counteroffer $(\hat{x}_2(b_1, 0), \hat{y}_2(b_1, 0))$ – player 2's payoff is $\delta\hat{y}_2(b_1, 0)$.

Using a technique similar to that used in deriving inequality 8, we have: $\delta\hat{y}_2((b_1, 0)) \leq \delta^2\dfrac{b_2}{2}$. Now since $\delta^2\gamma_2((b_1, 0), (0, b_2)) > \delta^2\dfrac{b_2}{2}$, we have $\delta^2\gamma_2((b_1, 0), (0, b_2)) > \delta\hat{y}_2((b_1, 0))$. Thus, player 2 must choose the second option, i.,e, player will reject $(b_1, 0)$ and offers $(0, b_2)$ at Stage 2.

2.8 Chapter 2 References

Anbarci, N. and J. H. Boyd, "Nash Demand Game and the Kalai-Smorodinsky Solution," *Games and Economic Behavior* 71 (2011), 14-22.

Dagan, N. and R. Serrano, "Invariance and Randomness in the Nash Program for Coalitional Games," *Economics Letters* 58 (1998), 43-49.

Deck, C. A. and A. Farmer, "Bargaining over an Uncertain Value: Arbitration Mechanisms Compared," *Journal of Law, Economics, and Organization* 23 (2007), 547-579.

Feuille, P., "Final Offer Arbitration and the Chilling Effect," *Industrial Relations* 14 (1975), 302-310.

Howard, J. V., "A Social Choice Rule and its Implementation in Perfect Equilibrium," *Journal of Economic Theory* 56 (1992), 142-159.

Kalai, E. and M. Smorodinsky, "Other Solutions to Nash's Bargaining Problem," *Econometrica* 43 (1975), 513-518.

Moulin, H., "Implementing the Kalai-Smorodinsky Bargaining Solution," *Journal of Economic Theory* 33 (1984), 32-45.

Nash, J., "The bargaining Problem," *Econometrica* 18 (1950), 155-162.

Nash, J., "Two-Person Cooperative Games," *Econometrica* 21 (1953), 128-140.

Rong, K., "Finite-Horizon Alternating-Offer Game with Arbitration," working paper, 2011.

Rubinstein, A., "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50 (1982), 97-109.

Serrano, R., "A Comment on the Nash Program and the Theory of Implementation," *Economics Letters* 55 (1997), 203-208.

Serrano, R., "Fifty Years of the Nash Program, 1953-2003," *Investigaciones Economicas* 29 (2005), 219-258.

Stevens, C., "Is Compulsory Arbitration Compatible with Bargaining?" *Industrial Relations* 5 (1966), 38-52.

Thomson, W., "Bargaining and the Theory of Cooperative Games: John Nash and beyond," working paper, 2009.

Yildiz, M., "Nash Meets Rubinstein in Final-Offer Arbitration," *Economics Letters* 110 (2011), 226-230.

Chapter 3

Proportional Individual Rationality and the Provision of a Public Good in a Large Economy

3.1 Introduction

The problem of public goods is a classic problem in economics. It is well-known that the voluntary provision of public goods may not attain the first-best efficient level due to the free-rider problem. How to design a budget-balanced mechanism that can achieve first-best efficiency is thus a fundamental topic in research in public economics. In the literature, if the (interim) individual rationality constraint is *not* required for any agent, then first-best efficiency can be achieved (d'Aspremont and Gérard-Varet 1979). If individual rationality is required for *all* agents, then first-best efficiency cannot be achieved (Mailath and Postlewaite 1990). This note considers a model that connects these two cases. Specifically, we consider a model in which the public good is provided and payments are collected from agents only if the proportion of agents who obtain nonnegative expected utilities from the mechanism weakly exceeds a prespecified ratio α for some constant $\alpha \in [0, 1]$. Roughly speaking, our model imposes individual rationality on a proportion α of agents. We call this requirement α *proportional individual rationality*. The parameter α is the *required* *agreement rate*. If α equals 1, then we return to the case of Mailath and Postlewaite (1990); if α equals 0, then we return to the case of d'Aspremont and Gérard-Varet (1979). Hence, both the model in d'Aspremont and Gérard-Varet (1979) and the model in Mailath and Postlewaite (1990) are special cases of our model.

We consider the public good provision problem in a *large* finite economy. We are not only interested in whether first-best efficiency can be achieved or not in a large economy, but also interested in the speed at which the efficiency or inefficiency is reached as the economy becomes large. Thus, two basic research questions of this note are as follows. First, for a given α, can the first-best provision level be achieved in a large economy? Second, how rapidly does the probability of provision approach its efficient or inefficient level as the economy becomes large?

We assume agents' valuations are i.i.d. and the total provision cost of the public good is proportional to the number of agents in the economy. Our results are as follows:

(i) If α is less than a threshold α^*, then there exists a sequence of mechanisms $\{\mu_\alpha^n\}_{n=1}^\infty$ such that for each n, μ_α^n satisfies budget balance, incentive compatibility, and α proportional individual rationality. As n goes to infinity, the probability of provision in μ_α^n approaches 1 at a speed not slower than $1/n$.

(ii) If α is greater than α^*, then for *any* sequence of anonymous mechanisms satisfying budget balance, incentive compatibility, and α proportional individual rationality, the probabilities of provision approach 0 at a speed not slower than $1/n^{\frac{1}{3}}$ as n goes to infinity.

The above results are summarized in Figure 31.

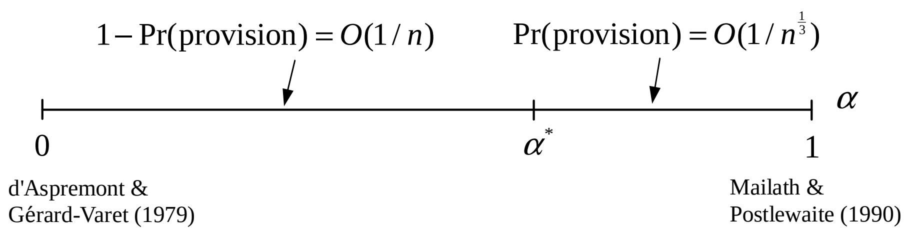

Figure 31: Probability of provision under various α as the economy becomes large.

The threshold α^* equals the probability of an agent's valuation being higher than the per capita cost of provision, i.e., $1 - F(c)$, where F is distribution function of an agent's valuation and c is the per capita cost of provision.

The implication of our results is as follows. Mailath and Postlewaite (1990) showed that if we require *unanimity*, then the public good will not be provided in a large economy. This implies that some coercion is necessary if the government wants the public good to be provided efficiently in a large economy. However, from the literature, it is not clear how much coercion is needed. Our results provide an answer for this question. In particular, we show that as soon as the required agreement rate α is less than $1 - F(c)$, the provision of the public good is asymptotically efficient. That is, if we want the public good to be provided efficiently in a large economy, then the proportion of agents that are allowed to be hurt in the mechanism should be *at least* $F(c)$. In another words, $F(c)$ is the *minimum* fraction of agents that the government must force to participate in order to obtain efficiency asymptotically.

This note is organized as follows. The next section presents the model. Section 3.1 discusses the asymptotic efficiency/inefficiency results. Section 3.2 discusses the convergence rates that efficiency/inefficiency obtains. Concluding remarks are offered in Section 3.4.

3.2 The Model

Assume that a nonexcludable public good can be provided in the quantity of either 0 unit or 1 unit. The cost of providing 1 unit of the public good is nc, where n is the number of agents in the economy, or the size of the economy, and c is a constant.[51] We denote agent i's valuation of the public goods by v_i. Agent i's valuation v_i is known only to agent i. We assume that $v_1, \cdots, v_n$ are independent and identically distributed.[52] The distribution function of v_i is denoted by F, which is common knowledge among all agents and the principal. The support of F is denoted by $[\underline{v}, \overline{v}] \subset R_+$. The density function of F is denoted by f. We assume that f is continuous and strictly positive on $[\underline{v}, \overline{v}]$. Finally, we assume that $\underline{v} < c < \overline{v}$.

We consider direct anonymous mechanisms. A direct mechanism is a function pair $\{q^n, \{t_i^n\}_{i=1}^n\}$ where $q^n : [\underline{v}, \overline{v}]^n \to \{0, 1\}$ indicates whether or not the public good is provided and $t_i^n : [\underline{v}, \overline{v}]^n \to R$ is the payment collected from agent i. Note that q^n and t_i^n are functions of reported valuations $v = (v_1, \cdots, v_n)$ of *all* agents. The anonymity of the mechanism requires q^n and t_i^n to be functions that depend exclusively on the reported valuations and not on the identities of agents.[53] Define $q_i^n(v_i) = E_{v_{-i}} q^n(v)$

[51]Here, the cost function is such that the cost of provision increases in proportion to the number of agents. This assumption is in the spirit of Mailath and Postlewaite (1990), in which the per capita cost of provision is bounded away from zero.

[52]We make the i.i.d. assumption for simplicity. Under this assumption, the characterization of the threshold α^* will be very simple. In the literature, the i.i.d. assumption also appeared in Rob (1989) and in Ledyard and Palfrey (2002).

[53]Formally, the anonymity of the mechanism requires (i) $q(v) = q(\sigma(v))$ for any permutation $\sigma(v)$ of v, (ii) $t_i^n(v_i, v_{-i}) = t_i^n(v_i, \sigma(v_{-i}))$ for any permutation $\sigma(v_{-i})$ of v_{-i}, and (iii) $t_i^n(v_i, v_{-i}) = t_j^n(v_j, v_{-j})$ where $v_i = v_j$ and v_{-j} is a permutation of v_{-i}. Anonymity of mechanisms is a key assumption for the results obtained in this note. In particular, if this assumption is dropped, then Theorem 3 in the next section will not hold, and we will obtain the efficiency result for *any* $\alpha \in [0, 1)$ in the large economy, because the principal can always pick an agent and charge him a tax equal to the total provision cost.

as agent i's *expected provision function* and $t_i^n(v_i) = E_{v_{-i}} t_i^n(v)$ as agent i's *expected payment function*. Given agents' reported valuations v, agent i's ex-post utility under the mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$ is given by $v_i q^n(v) - t_i^n(v)$. Define $U_i^n(v_i, \hat{v}_i) = E_{v_{-i}}[v_i q^n(\hat{v}_i, v_{-i}) - t_i^n(\hat{v}_i, v_{-i})]$ and $U_i^n(v_i) = U_i^n(v_i, v_i)$. Then $U_i^n(v_i, \hat{v}_i)$ represents agent i's (interim) expected utility when he reports $\hat{v}_i$ and $U_i^n(v_i)$ represents agent i's (interim) expected utility when he reports truthfully, both conditional on all other agents reporting truthfully.

We impose three constraints on the mechanism. The first constraint is the (ex ante) *budget balance* constraint (BB). The second constraint is the (interim) *incentive compatibility* constraint (IC).

$$(\text{BB}) : E\{\frac{\sum t_i^n(v)}{n} - cq^n(v)\} \geq 0 \tag{9}$$

$$(\text{IC}) : U^n(v_i) \geq U^n(v_i, \hat{v}_i) \text{ for all } v_i, \hat{v}_i \in [\underline{v}, \bar{v}] \tag{10}$$

The last constraint is the α *proportional individual rationality* constraint. This constraint reflects the requirement that the public good can be provided and payments can be collected from agents only if the proportion of agents who obtain nonnegative expected utilities from the mechanism is at least α. Define the *agreement set* $\tilde{V}_i^n$ of agent i as the set of agent i's valuations for which the expected utility of agent i is nonnegative, i.e., $\tilde{V}_i^n = \{v_i \in [\underline{v}, \bar{v}] | U_i^n(v_i) \geq 0\}$. Define $r^n(v) = \frac{\sum_{i=1}^n \mathbf{1}_{\{v_i \in \tilde{V}_i^n\}}}{n}$ where $\mathbf{1}_{\{\bullet\}}$ represents the indicator function. Then, $r^n(v)$ is the proportion of agents who obtain nonnegative expected utilities from the mechanism. The α *proportional individual rationality* constraint is:

$$(\alpha\text{-PIR}) : q^n(v) = 0 \text{ and } t_i^n(v) = 0 \text{ (for all } i\text{) if } r^n(v) < \alpha.$$

In this note, α is exogenously fixed and can be any number between 0 and 1. If α=0, then the α proportional individual rationality constraint is automatically satisfied by any mechanism. The following lemma establishes the connection between α proportional individual rationality and individual rationality (IR).

Lemma 2. *A mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$ satisfies (1-PIR) if and only if it satisfies (IR).*

Proof: The "if" part is obvious. To see the "only if" part, suppose there is a $v_i^* \in [\underline{v}, \overline{v}]$ such that $U_i^n(v_i^*) < 0$, then we must have $r^n(v_i^*, v_{-i}) < 1$ for all v_{-i}, which implies $q^n(v_i^*, v_{-i}) = 0$ and $t_i^n(v_i^*, v_{-i}) = 0$ for all v_{-i} by the α proportional individual rationality constraint. Thus, $U_i(v_i^*)$ must equal zero, which is a contradiction with $U_i(v_i^*) < 0$. □

We now define the first-best mechanism. The first-best mechanism $\{q^{FB(n)}, \{t_i^{FB(n)}\}_{i=1}^n\}$ is any budget-balanced mechanism that satisfies the Lindahl-Samuelson provision rule:

$$q^{FB(n)} = \begin{cases} 1 & \text{if } \dfrac{\sum v_i}{n} \geq c; \\ 0 & \text{otherwise.} \end{cases}$$

We assume that $v^e > c$, where $v^e = E(v_i)$. This assumption implies that in a large economy, the average social benefit of the public good is greater than the average social cost of the public good. As a result, the public good should be provided with probability 1 in the large economy, i.e., $P(q^{FB(n)}(v) = 1) \to 1$ as $n \to \infty$.

3.3 Analysis

3.3.1 Asymptotic efficiency/inefficiency results

Let $\mathcal{M}_\alpha^n = \{\{q^n, \{t_i^n\}_{i=1}^n\} | \{q^n, \{t_i^n\}_{i=1}^n\}$ is anonymous and satisfies BB, IC, and α-PIR$\}$ be the set of *feasible* mechanisms in the n-agent economy, and let $s_\alpha^n = \sup\{P(q^n(v) = 1) : \exists \{q^n, \{t_i^n\}_{i=1}^n\} \in \mathcal{M}_\alpha^n\}$ be the maximum probability of provision in the set of feasible mechanisms of the n-agent economy. Let $\alpha^* = 1 - F(c)$. We have:

Theorem 3.

*(i) If $\alpha < \alpha^*$, then s_α^n goes to **one** as $n \to \infty$;*

*(ii) If $\alpha > \alpha^*$, then s_α^n goes to **zero** as $n \to \infty$.*

Proof: Theorem 3 follows from Theorem 4 and Theorem 5 in the next subsection. The proof below is a simple intuitive proof.

When n is large, the probability that an agent is pivotal in *any* anonymous mechanism is small. By the incentive compatibility constraint, an agent's expected payment function (i.e., $t_i^n(v_i)$) must be nearly constant, independent of the agent's valuation.[54] By the budget balance constraint, the constant tax t on each agent

[54] The reason is as follows. Using the incentive compatibility constraint and the fact that $[\underline{v}, \overline{v}] \subset R_+$, one can show that if an agent reports a low valuation, then he pays a small tax. If he reports a high valuation, then he pays a large tax. This difference in tax is proportional to the probability that the agent is pivotal in the mechanism. As the economy becomes large, the probability that an agent is pivotal in any mechanism becomes small. This means that the agent with low valuation and the agent with high valuation must pay almost the same tax in a large economy.

must be equal to the per capita cost c. This implies that the proportion of agents who obtain nonnegative expected utilities in a large economy must be approximately $P(v_i \geq c) = 1 - F(c)$. By the α-proportional individual rationality constraint, whether or not the public good will be provided in the large economy then depends on whether or not $\alpha \leq 1 - F(c)$. □

Theorem 3 states that if α is small, then we have the efficiency result, and if α is large, then we have the inefficiency result. These results are not surprising considering that we already have the two extreme results in the literature (i.e., the case $\alpha = 0$ and the case $\alpha = 1$). However, Theorem 3 provides us with the *exact* threshold that we need. In addition, it shows that the probability of provision is either 0 or 1 in the large economy, i.e., the probability of provision is "discontinuous" in the required agreement rate α.

3.3.2 Convergence rates

This section will analyze the convergence rates at which efficiency or inefficiency is attained as the economy becomes large.

The case where $\alpha < 1 - F(c)$

In this case, we will construct a particular mechanism, called the α-*referendum*, which attains first-best efficiency asymptotically, and we will find the convergence rate of the probability of provision in the α-referendum. The α-referendum, which is denoted by $\mu^n_\alpha = \{\tilde{q}^{n,\alpha}, \{\tilde{t}^{n,\alpha}_i\}^n_{i=1}\}$, is constructed as follows. If $\dfrac{\sum_{i=1}^n \mathbf{1}_{\{v_i \geq c\}}}{n} \geq \alpha$, then $\tilde{q}^{n,\alpha}(v) = 1$ and $\tilde{t}^{n,\alpha}_i(v) = c$ for any i. If $\dfrac{\sum_{i=1}^n \mathbf{1}_{\{v_i \geq c\}}}{n} < \alpha$, then $\tilde{q}^{n,\alpha}(v) = 0$ and $\tilde{t}^{n,\alpha}_i(v) = 0$. Simply speaking, μ^n_α requires that if there exists at least α proportion of agents whose reported valuations of the public good exceed c, then the public good will be provided and the cost of provision will be distributed equally among all agents. Otherwise, the public good will not be provided and no payment is collected from any agent.

One can easily verify that μ_α^n satisfies (BB), (IC), and α-(PIR).[55] We have the following convergence rate result for the α-referendum.

Theorem 4. *If $\alpha < 1 - F(c)$, then $1 - P(\tilde{q}^{n,\alpha}(v) = 1) = O(1/n)$.*

Proof: See Appendix 1. □

The probability of provision in μ_α^n converges to 1 as the economy becomes large, because the proportion of agents whose valuations exceed c approaches $P(v_i \geq c) = 1 - F(c)$, which is greater than α by assumption. Using Chebyshev's inequality, one can show that the convergence speed of $P(\tilde{q}^{n,\alpha}(v) = 1)$ toward 1 is on the order of $1/n$.

The performance of α-referenda is illustrated in the following example.

Example 1: Assume that v_i is uniformly distributed on $[0, 1]$. Assume $c = 0.4$ and $\alpha = 0.5$.

Table 4 lists the probabilities that the public good will be provided in 0.5-referenda as n increases from 4 to 200. It also lists the per capita welfare of 0.5-referenda,[56] the per capita welfare of first-best mechanisms and the per capita welfare

[55] μ_α^n satisfies α-PIR because under μ_α^n, (i) an agent obtains a nonnegative expected utility if and only if the agent's valuation exceeds c, and (ii) the public good will be provided and payments will be collected from agents only if the proportion of agents whose valuations exceed c is greater than the required agreement rate α.

[56] The per capita welfare of a mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$ is defined as the per capita expected value

losses from using 0.5-referenda instead of first-best mechanisms.

Table 4: Probability of provision and relative efficiency of 0.5-referenda

n	$1-P(\tilde{q}^{n,0.5}=1)$	$W(\tilde{q}^{n,0.5})$	$W(q^{FB(n)})$	$1-\frac{W(\tilde{q}^{n,0.5})}{W(q^{FB(n)})}$
5	0.3169	0.1097	0.1165	0.0581
10	0.1662	0.1034	0.1062	0.0266
20	0.1273	0.0990	0.1016	0.0264
40	0.0747	0.0981	0.1002	0.0209
80	0.0267	0.0990	0.1001	0.0095
160	0.0041	0.0997	0.1000	0.0021
320	0.0001	0.0999	0.0999	0.0001

Table 4 shows that as n increases, $1-P(\tilde{q}^{n,0.5})$ approaches 0 rapidly. Although the convergence speed is slower than $1/n$ when n is small ($n \leq 40$), it is eventually faster than $1/n$ when n becomes large ($n > 40$). Table 4 also shows that the per capita welfare loss from the 0.5-referendum vanishes rapidly as n increases. When the size of the economy is greater than 80, the per capita welfare loss is less than 1%.

The case where $\alpha > 1-F(c)$

In the case where the required agreement rate α exceeds the threshold $1-F(c)$, we have the inefficiency result. In this case, instead of finding the convergence rate of a particular mechanism, we will find the convergence rate that holds for *all* feasible mechanisms. In particular, we have the following result.[57]

of the net benefit of the mechanism, i.e., $W(q^n) = \frac{1}{n}E\left\{(\sum v_i - nc)q^n(v)\right\}$.

[57]For the special case where $\alpha = 1$, the convergence rate we obtain in Theorem 5 is the same as the convergence rate obtained by Al-Najjar and Smorodinsky (2000). The difference between the two results is that Al-Najjar and Smorodinsky (2000) assumed that the distribution of an agent's valuation is *discrete* at $\underline{v}$. They need this assumption because their characterization about a player's influence in a mechanism is more suitable for the case where the distribution of an agent's valuation is discrete, while we do not need this assumption because our characterization about a

Theorem 5. *For any given $\alpha > 1 - F(c)$, we have*

$$s_\alpha^n = O(1/n^{\frac{1}{3}}).$$

Proof: See Appendix 2. $\square$

The intuition of the proof of Theorem 5 is as follows. The fact that as the economy becomes large, the probability that an agent is pivotal in any mechanism becomes small implies that the expected provision function of any agent must be nearly constant in a large economy. The convergence speed at which s_α^n goes to zero depends on the convergence speed at which the expected provision function of an agent approaches the constant function. In Lemma 7 (Appendix 2), we show that the expected provision function of any agent approaches a constant function at a speed of $1/\sqrt{n}$. Based on this result, one can show that the probability of provision $P(q^n(v) = 1) \leq \dfrac{C_\alpha}{\sqrt{\eta n}} + \eta$ for any sufficiently small $\eta > 0$. This implies that s_α^n goes to zero at a speed of $1/n^{\frac{1}{3}}$.

The following example illustrates the analytic bound obtained in the proof of Theorem 5.

***Example 2*:** Assume v_i follows uniform distribution on $[0, 1]$. Assume $c = 0.4$, $\bar{t} = 1$ and $\epsilon = 0.01$.[58] Suppose α can take values of 0.67 and 1.

The threshold $1 - F(c) = 0.6$ is less than the required agreement rates 0.67 and 1. Figure 32 illustrates the upper bounds for the probabilities of provision as n varies

player's influence in a mechanism is more suitable for the case where the distribution of an agent's valuation is continuous (see also footnote 59 in Appendix 2).

[58] The parameter ϵ can be any number between 0 and $c - v_\alpha$, where v_α is such that $P(v_i \geq v_\alpha) = \alpha$. The precise choice of ϵ will not affect the convergence rate.

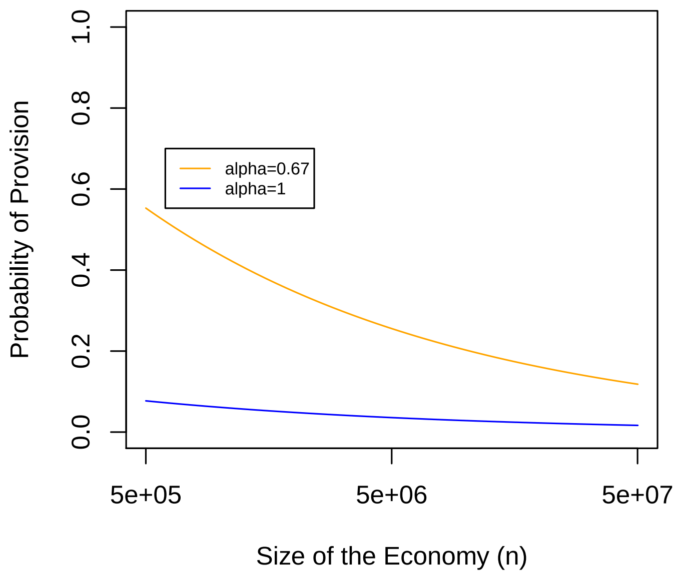

Figure 32: Upper bounds for the probability of provision when $\alpha > 1 - F(c)$.

from 50,000 to 50,000,000. For both the case of $\alpha = 0.67$ and the case of $\alpha = 1$, the upper bound for the probability of provision decreases to 0 as n increases. In particular, when $n = 50,000$, the corresponding upper bounds for the probability of provision are 0.553 for $\alpha = 0.67$ and 0.077 for $\alpha = 1$. When $n = 50,000,000$, the bounds are 0.11 for $\alpha = 0.67$ and 0.017 for $\alpha = 1$. Obviously, the bound for the probability of provision for the case of $\alpha = 0.67$ is consistently higher than the bound for the probability of provision for the case of $\alpha = 1$. This is not surprising because for any mechanism that satisfies (α_1-PIR), it must also satisfy (α_2-PIR) if $\alpha_2 < \alpha_1$. Thus, as α decreases, the set of mechanisms that satisfy the α-PIR constraint becomes larger and the upper bound for the probability of provision is higher.

3.4 Conclusion

This note considers the public good provision problem in which the public good can be provided and payments can be collected from agents only if the proportion of agents who obtain nonnegative expected utilities from the mechanism weakly exceeds a required agreement rate α. We show that, if α is less than a threshold, then there exists a sequence of α-referenda satisfying BB, IC and α-PIR such that as the size n of the economy becomes large, the probabilities of the public good being provided in α-referenda approach 1 at a speed not slower than $1/n$; and if α exceeds the threshold, then the probabilities that the public good will be provided in any series of anonymous mechanisms that satisfy BB, IC, and α-PIR approach 0 at a speed not slower than $1/n^{\frac{1}{3}}$. Hence, we not only obtain asymptotic efficiency/inefficiency results for various required agreement rates settings, but also characterize the speed at which the probability of provision reaches the efficient/inefficient level.

3.5 Chapter 3 Appendix

3.5.1 Appendix 1: Proof of Theorem 4

Proof: Noting that $\frac{\sum_{i=1}^{n}\mathbf{1}_{\{v_i\geq c\}}}{n} \to 1-F(c)$ in pr. and $Var(\mathbf{1}_{\{v_i\geq c\}}) = (1-F(c))F(c)$, we have:

$$\begin{aligned} P(\tilde{q}^{n,\alpha}=1) = P\Big(\frac{\sum_{i=1}^{n}\mathbf{1}_{\{v_i\geq c\}}}{n}\geq\alpha\Big) \quad &\geq \quad P\left(\left|1-F(c)-\frac{\sum_{i=1}^{n}\mathbf{1}_{\{v_i\geq c\}}}{n}\right|\leq 1-F(c)-\alpha\right) \\ &\geq \quad 1-\frac{(1-F(c))F(c)}{n(1-F(c)-\alpha)^2}. \end{aligned}$$

where the last inequality follows from Chebyshev's inequality. Thus, $P(\tilde{q}^{n,\alpha} = 1) \geq 1 - O(1/n)$. $\square$

3.5.2 Appendix 2: Proof of Theorem 5

We use the following three lemmas (Lemma 6, Lemma 7 and Lemma 8) to prove Theorem 5.

Lemma 6 states that the variance of $q_i^n(v_i)$ converges to 0 at a speed not slower than $1/n$ as $n \to \infty$.

Lemma 6. *For any anonymous mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$, we have $Var(q_i^n(v_i)) \leq \dfrac{Var(q^n(v))}{n}$. Thus, $Var(q_i^n(v_i)) = O(1/n)$.*

Proof: Let $V = [\underline{v}, \overline{v}]^n$ and $V_i = [\underline{v}, \overline{v}]$ for any i. We have:

$$\begin{aligned}
&\int_V [q^n(v) - Eq^n(v) - \sum_i (q_i^n(v_i) - Eq_i^n(v_i))]^2 dF(v_1)\cdots F(v_n) \\
= &\int_V [q^n(v) - Eq^n(v)]^2 dF(v_1)\cdots F(v_n) \\
&- 2\sum_i \int_V (q^n(v) - Eq^n(v))(q_i^n(v_i) - Eq_i^n(v_i)) dF(v_1)\cdots F(v_n) \\
&+ \sum_i \int_V [q_i^n(v_i) - Eq_i^n(v_i)]^2 dF(v_1)\cdots F(v_n)
\end{aligned}$$

$$
\begin{aligned}
&= \int_V [q^n(v) - Eq^n(v)]^2 dF(v_1)\cdots F(v_n) - 2\sum_i \int_{V_i} [q_i^n(v_i) - Eq_i^n(v_i)]^2 dF(v_i) \\
&\quad + \sum_i \int_{V_i} [q_i^n(v_i) - Eq_i^n(v_i)]^2 dF(v_i) \\
&= Var(q^n(v)) - \sum_i Var(q_i^n(v_i)) \\
&= Var(q^n(v)) - nVar(q_i^n(v_i)) \qquad (11)
\end{aligned}
$$

where the first equality follows from the fact that $q_1^n(v_1), \cdots, q_n^n(v_n)$ are independent, and the second equality follows from the fact that $Eq^n(v) = Eq_i^n(v_i)$ for any i.

Using equality (11) and the fact that $\int_V [q^n(v) - Eq^n(v) - \sum_i (q_i^n(v_i) - Eq_i^n(v_i))]^2 dF(v_1)\cdots F(v_n) \geq 0$, we have $Var(q_i^n(v_i)) \leq \frac{1}{n} Var(q^n(v))$. Notice that $0 \leq q^n(v) \leq 1$ for any n and $v \in V$, so $Var(q^n(v)) \leq 1$. Thus, $Var(q_i^n(v_i)) = O(1/n)$. $\square$

An agent i's *influence* relative to the mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$ can be defined by $\sqrt{Var(q_i^n(v_i))}$. Lemma 6 says that as n goes to infinity, an agent's influence in any sequence of mechanisms decreases to zero at a speed not slower than $1/\sqrt{n}$.[59]

The next lemma is a direct result of Lemma 6. It says that for any $\tilde{v} \in (\underline{v}, \bar{v})$, the difference between $q_i^n(\tilde{v})$ and $Eq_i^n(v_i)$ vanishes as $n \to \infty$ and is $O(1/\sqrt{n})$.

For any given $\tilde{v} \in (\underline{v}, \bar{v})$, define $m(\tilde{v}) = \min(P(v_i \geq \tilde{v}), P(v_i \leq \tilde{v}))$ (notice that $m(\tilde{v}) > 0$ because the probability density function of v_i is positive on $[\underline{v}, \overline{v}]$). We

[59]This result is similar to the result obtained by Al-Najjar and Smorodinsky (2000), who use the difference between the maximum and the minimum of the expected provision function to measure agent i's influence. They showed that in any mechanism, the average influence of agents is $O(1/\sqrt{n})$. While the measure of influence in Al-Najjar and Smorodinsky (2000) is more suitable for the case where the distributions of valuations are *discrete*, our measure of influence is more suitable for the case where the distributions of valuations are *continuous*.

have:

Lemma 7. *Let $\{\{q^n, \{t_i^n\}_{i=1}^n\}\}_{n=1}^{\infty}$ be any sequence of anonymous mechanisms where, for each n, $\{q^n, \{t_i^n\}_{i=1}^n\}$ satisfies (IC). Then, for any given $\tilde{v} \in (\underline{v}, \bar{v})$, we have:*

$$|q_i^n(\tilde{v}) - Eq_i^n(v_i)| \leq \frac{1}{2\sqrt{m(\tilde{v})n}}.$$

Proof: First, notice that if $\{q^n, \{t_i^n\}_{i=1}^n\}$ satisfies (IC), then $q_i^n(v_i)$ is nondecreasing on $[\underline{v}, \bar{v}]$.

For a given $\varepsilon > 0$, using Chebyshev's inequality, Lemma 6 and the fact that $Var(q^n(v)) \leq \dfrac{1}{4}$ (note $q^n(v)$ follows a Bernoulli distribution that can take two values: 0 and 1), we have:

$$P(|q_i^n(v_i) - Eq_i^n(v_i)| \geq \varepsilon) \leq \frac{Var(q_i^n(v_i))}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

Fix $\tilde{v} \in (\underline{v}, \bar{v})$. If $n \geq \dfrac{1}{4m(\tilde{v})\varepsilon^2}$, then

$$P(|q_i^n(v_i) - Eq_i^n(v_i)| > \varepsilon + \varepsilon_0) \leq \frac{1}{4n(\varepsilon+\varepsilon_0)^2} < m(\tilde{v}) \quad \text{for any } \varepsilon_0 > 0. \tag{12}$$

Now if $|q_i^n(\tilde{v}) - Eq_i^n(v_i)| > \varepsilon + \varepsilon_0$, then we have either $q_i^n(\tilde{v}) - Eq_i^n(v_i) > \varepsilon + \varepsilon_0$ or $q_i^n(\tilde{v}) - Eq_i^n(v_i) < -(\varepsilon + \varepsilon_0)$. Since $q_i^n(v_i)$ is an nondecreasing function, we have either $q_i^n(v_i) - Eq_i^n(v_i) > (\varepsilon + \varepsilon_0)$ for all $v_i \geq \tilde{v}$ or $q_i^n(v_i) - Eq_i^n(v_i) < -(\varepsilon + \varepsilon_0)$ for all $v_i \leq \tilde{v}$. Thus, $P(|q_i^n(v_i) - Eq_i^n(v_i)| > \varepsilon + \varepsilon_0) \geq \min(P(v_i \geq \tilde{v}), P(v_i \leq \tilde{v})) = m(\tilde{v})$ which is a contradiction with (12). As a result, we must have $|q_i^n(\tilde{v}) - Eq_i^n(v_i)| \leq \varepsilon + \varepsilon_0$ for all $n \geq \dfrac{1}{4m(\tilde{v})\varepsilon^2}$. Since $\varepsilon_0 > 0$ is arbitrary, then $|q_i^n(\tilde{v}) - Eq_i^n(v_i)| \leq \varepsilon$ for all

$n \geq \dfrac{1}{4m(\tilde{v})\varepsilon^2}$. Hence, for any given N, we have $|q_i^n(\tilde{v}) - Eq_i^n(v_i)| \leq \dfrac{1}{2\sqrt{m(\tilde{v})N}}$ for all $n \geq N$. $\square$

In order to obtain the convergence speed that the expected payment function goes to the constant function, we need the assumption that $|t_i^n(v)|$ is uniformly bounded in n and v (Assumption 1). We assume that Assumption 1 holds through the remainder of the appendix. Notice that it is a natural assumption to put a bound on the payment that an individual agent can be forced to contribute (see, e.g., Mailath and Postlewaite 1990).

Assumption 1: $|t_i^n(v)|$ is uniformly bounded, i.e., there exists a $\bar{t} > 0$ such that $|t_i^n(v)| \leq \bar{t}$ for any n and $v \in [\underline{v}, \overline{v}]^n$.

For any $\eta \in (0, \frac{1}{2})$, define v_η as the only $v_i \in [\underline{v}, \overline{v}]$ such that $F(v_\eta) = \eta$ and v'_η as the only $v_i \in [\underline{v}, \overline{v}]$ such that $1 - F(v'_\eta) = \eta$.

Lemma 8. *Let $\{\{q^n, \{t_i^n\}_{i=1}^n\}\}_{n=1}^\infty$ be any sequence of anonymous mechanisms where, for each n, $\{q^n, \{t_i^n\}_{i=1}^n\}$ satisfies (IC). Then, for any $\eta \in (0, \frac{1}{2})$ and $\tilde{v} \in (v_\eta, v'_\eta)$, we have: $|t_i^n(\tilde{v}) - Et_i^n(v_i)| \leq \overline{v}\dfrac{1}{\sqrt{\eta n}} + 2\eta\bar{t}$.*

Proof: Fix $\tilde{v} \in (v_\eta, v'_\eta)$. For any $v_i \in (v_\eta, v'_\eta)$, we have $t_i^n(v_i) \leq v_i(q_i^n(v_i) - q_i^n(\tilde{v}_i)) + t_i^n(\tilde{v}) \leq \overline{v}\dfrac{1}{\sqrt{\eta n}} + t_i^n(\tilde{v})$, where the first inequality follows from the incentive compatibility constraint and the second inequality follows from Lemma 7. Thus, $Et_i^n(v_i) = \int_{(v_\eta, v'_\eta)} t_i^n(v_i)dF_i(v_i) + \int_{[\underline{v}, v_\eta]} t_i^n(v_i)dF_i(v_i) + \int_{[v'_\eta, \overline{v}]} t_i^n(v_i)dF_i(v_i) \leq \overline{v}\dfrac{1}{\sqrt{\eta n}} + t_i^n(\tilde{v}) + 2\eta\bar{t}$. That is, $t_i^n(\tilde{v}) - Et_i^n(v_i) \geq -\overline{v}\dfrac{1}{\sqrt{\eta n}} - 2\eta\bar{t}$. Similarly, it can be shown that $t_i^n(\tilde{v}) - Et_i^n(v_i) \leq \overline{v}\dfrac{1}{\sqrt{\eta n}} + 2\eta\bar{t}$. Thus, $|t_i^n(\tilde{v}) - Et_i^n(v_i)| \leq \overline{v}\dfrac{1}{\sqrt{\eta n}} + 2\eta\bar{t}$. $\square$

Since the prior distribution of v_i is the same for all i and the mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$ is anonymous by assumption, the function $U_i^n(v_i)$ must be the same for all i. In the following, we will use $U^n(v_i)$ to denote $U_i^n(v_i)$, whenever there is no confusion.

It is well-known in the mechanism design literature that if the mechanism $\{q^n, \{t_i^n\}_{i=1}^n\}$ satisfies incentive compatibility and q^n is bounded, then the function $U^n(v_i)$ is continuous. This implies that $U^n(v_i)$ is either (i) equal to zero at some point in $[\underline{v}, \overline{v}]$, (ii) greater than zero for all $v_i \in [\underline{v}, \overline{v}]$, or (iii) less than zero for all $v_i \in [\underline{v}, \overline{v}]$. We define $\hat{v}^n$ as follows:

$$\hat{v}^n = \begin{cases} \min\{v_i | v_i \in [\underline{v}, \bar{v}] \text{ and } U^n(v_i) = 0\} & if \quad U^n(v_i) = 0 \text{ for some } v_i \in [\underline{v}, \overline{v}]; \\ \underline{v} & if \quad U^n(v_i) > 0 \text{ for all } v_i \in [\underline{v}, \overline{v}]; \\ \infty & if \quad U^n(v_i) < 0 \text{ for all } v_i \in [\underline{v}, \overline{v}]. \end{cases}$$

Using the definition of $\hat{v}^n$ and the fact that $U^n(v_i)$ is nondecreasing in v_i (by incentive compatibility), we have

$$\tilde{V}_i^n = \{v_i \in [\underline{v}, \bar{v}] | U^n(v_i) \geq 0\} = \{v_i \in [\underline{v}, \bar{v}] | v_i \geq \hat{v}^n\}.$$

For a given α, define $v_\alpha \in [\underline{v}, \bar{v}]$ such that $P(v_i \geq v_\alpha) = \alpha$. Note that $\alpha > 1 - F(c)$ implies $v_\alpha < c$.

Now, we can state the proof of Theorem 5.

Proof of Theorem 5:

For any given n and any given $\epsilon \in (0, c - v_\alpha)$, we have two cases.

(i) $\hat{v}^n < v_\alpha + \epsilon$.

In this case, since $\hat{v}^n < v_\alpha + \epsilon < \bar{v}$, we must have $U^n(\hat{v}^n) \geq 0$ by the definition of $\hat{v}^n$. The incentive compatibility constraint implies that $U^n(v_i)$ is nondecreasing in v_i, so $U^n(v_\alpha + \epsilon) \geq U^n(\hat{v}^n) \geq 0$. That is:

$$(v_\alpha + \epsilon) q_i^n(v_\alpha + \epsilon) \geq t_i^n(v_\alpha + \epsilon). \tag{13}$$

Since $\{q_i^n(v_i), t_i^n(v_i)\}$ satisfies (IC), then by Lemma 7, we have:

$$q_i^n(v_\alpha + \epsilon) \leq Eq_i^n(v_i) + \frac{1}{2\sqrt{m(v_\alpha + \epsilon)n}} \tag{14}$$

By Lemma 8, for sufficiently small η, we have:

$$t_i^n(v_\alpha + \epsilon) \geq Et_i^n(v_i) - \overline{v}\frac{1}{\sqrt{\eta n}} - 2\eta\bar{t} \tag{15}$$

(13) (14) and (15) then imply an upper bound for $Et_i^n(v_i)$, that is:

$$Et_i^n(v_i) \leq (v_\alpha + \epsilon)Eq_i^n(v_i) + \frac{v_\alpha + \epsilon}{2\sqrt{m(v_\alpha + \epsilon)n}} + \overline{v}\frac{1}{\sqrt{\eta n}} + 2\eta\bar{t} \tag{16}$$

Now, by the budget balance constraint, we can get a lower bound for $Et_i^n(v_i)$, that is:

$$Et_i^n(v_i) \geq cEq_i^n(v_i). \tag{17}$$

Inequality (16) and inequality (17) then imply:

$$Eq_i^n(v_i) \leq \frac{1}{c - (v_\alpha + \epsilon)}\left(\frac{v_\alpha + \epsilon}{2\sqrt{a(v_\alpha + \epsilon)}}\frac{1}{\sqrt{n}} + \overline{v}\frac{1}{\sqrt{\eta}}\frac{1}{\sqrt{n}} + 2\eta\bar{t}\right). \tag{18}$$

Since $Eq_i^n(v_i) = Eq^n(v) = P(q^n(v) = 1)$, then we have:

$$P(q^n(v) = 1) \leq \frac{1}{c - (v_\alpha + \epsilon)}\left(\frac{v_\alpha + \epsilon}{2\sqrt{a(v_\alpha + \epsilon)}}\frac{1}{\sqrt{n}} + \overline{v}\frac{1}{\sqrt{\eta}}\frac{1}{\sqrt{n}} + 2\eta\bar{t}\right). \tag{19}$$

The bound obtained in (19) implies that $P(q^n(v) = 1) = O(1/n^{\frac{1}{3}})$ for any sequence of mechanisms $\{\{q^n, \{t_i^n\}_{i=1}^n\}\}_{n=1}^\infty$ where $\{q^n, \{t_i^n\}_{i=1}^n\} \in \mathcal{M}_\alpha^n$ for each n and $\alpha > 1 - F(c)$.[60]

(ii) $\hat{v}^n \geq v_\alpha + \epsilon$.

Define $\gamma^n(v) = \dfrac{\sum_{i=1}^n \mathbf{1}_{\{v_i \geq v_\alpha + \epsilon\}}}{n}$. Note $r^n(v) = \dfrac{\sum_{i=1}^n \mathbf{1}_{\{v_i \geq \hat{v}^n\}}}{n}$, then we must have $\gamma^n \geq r^n$. By (α-PIR), if $r^n < \alpha$ then $q^n(v) = 0$. Thus, we have $P(q^n(v) = 1) \leq P(r^n \geq \alpha) \leq P(\gamma^n \geq \alpha)$. Since $\gamma^n \to P(v_i \geq v_\alpha + \epsilon) := \beta_0$ in pr. by the law of large numbers, we have $P(\gamma^n \geq \alpha) = P(\gamma^n - \beta_0 \geq \alpha - \beta_0) \leq P(|\gamma^n - \beta_0| \geq \alpha - \beta_0) \leq \dfrac{Var(\gamma^n)}{n(\alpha - \beta_0)^2} = \dfrac{\beta_0(1-\beta_0)}{n(\alpha - \beta_0)^2}$, where the second inequality follows from Chebyshev's inequality. Hence,

$$P(q^n(v) = 1) \leq \frac{\beta_0(1-\beta_0)}{(\alpha - \beta_0)^2}\frac{1}{n}. \tag{20}$$

For any given n, the probability of provision $P(q^n(v) = 1)$ is either bounded by inequality (19), or bounded by inequality (20). This implies that $P(q^n(v) = 1) = O(1/n^{\frac{1}{3}})$ for any sequence of mechanisms $\{\{q^n, \{t_i^n\}_{i=1}^n\}\}_{n=1}^\infty$ where $\{q^n, \{t_i^n\}_{i=1}^n\} \in \mathcal{M}_\alpha^n$ for each n and $\alpha > 1 - F(c)$. $\square$

[60]Notice that the bound obtained in (19) is the tightest when $\eta = (\frac{4\bar{t}}{\overline{v}})^{-\frac{2}{3}} n^{-\frac{1}{3}}$, which implies that $P(q^n(v) = 1) = O(1/n^{\frac{1}{3}})$.

3.6 Chapter 3 References

N. I. Al-Najjar, R. Smorodinsky, *Pivotal players and the characterization of influence*, Journal of Economic Theory 92 (2000), 318-342.

C. d'Aspremont, L. A. Gérard-Varet, *Incentives and incomplete information*, Journal of Public Economics 11 (1979), 25-45.

J. O. Ledyard, *Non-Coercion, efficiency and incentive compatibility in public goods decisions*, working paper, 2010.

J. O. Ledyard, T. R. Palfrey, *The approximation of efficient public good mechanisms by simple voting schemes*, Journal of Public Economics 83 (2002), 153-171.

G. Mailath, A. Postlewaite, *Asymmetric information bargaining problems with many agents*, Review of Economic Studies 57 (1990), 531-567.

R. Rob, *Pollution claim settlements under private information*, Journal of Economic Theory 47 (1989), 307-333.